



MASTERARBEIT | MASTER'S THESIS

Titel | Title

Differentiation of 24 Chickpea Cultivars by Targeting
Microsatellites and Single Nucleotide Polymorphisms using
High Resolution Melting Analysis

verfasst von | submitted by
Isabel Niederkofler BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree
programme code as it appears on the
student record sheet:

UA 066 659

Studienrichtung lt. Studienblatt | Degree
programme as it appears on the student
record sheet:

Masterstudium Lebensmittelchemie

Betreut von | Supervisor:

Univ.-Prof. Mag. Dr. Margit Cichna-Markl

Acknowledgments

First and foremost, I would like to thank my supervisor, Univ.-Prof. Mag. Dr. Margit Cichna-Markl for the exceptional guidance during my thesis. I enjoyed the time I spent in her lab and learned a lot during this period.

I would also like to express my gratitude to Jakob Scheibelreiter MSc, for showing me around and Katharina Pühringer MSc, for the helpful tips and good conversations.

I would also like to thank our cooperation partners at AGES, Dipl.-Ing.in Magdalena Wagner, Dr. Elisabeth Reiter and Dr. Stefano D'Amico, for providing the chickpea samples and Nathalie Runge MSc for grinding the samples for me.

I am deeply grateful for my fiance Daniel, who introduced me to LaTeX, which was used to write this thesis, and for the emotional support throughout my studies.

Thank you to my friends - those I met during my Master's and those who were already there before - for always being there for me when I need them most. I would also like to thank my parents for their financial support during my studies.

Lastly, I would like to thank the current team of the WG Cichna, including Benno Fehringer BSc, Klementina Yakova BSc and Sofie Schandl BSc, for the entertaining hours spent together in the lab.

I dedicate this thesis to my late father, who was very proud of me starting my thesis but did not live to see it finished.

Abstract

Cultivation of chickpeas is becoming more interesting with the progressing climate change, as chickpeas grow better in a warmer environment and are less sensitive to water limitations. Additionally, chickpeas provide a good source of protein and soluble and insoluble fibers, show a low content of fat and sodium, and contain no cholesterol. With the rising trend of vegan products and meat alternatives, food containing chickpeas becomes more popular in European supermarkets.

This study aimed to differentiate 24 chickpea cultivars belonging to the three types desi, kabuli and gulabi, most of them originating from Europe. Apart from the visual differences between desi, kabuli and gulabi, no morphological distinction is possible among cultivars belonging to one of these three types. We developed DNA-based assays including polymerase chain reaction (PCR) to amplify chickpea DNA containing microsatellites or single nucleotide polymorphisms (SNPs), followed by high resolution melting (HRM) analysis.

Multiple primer pairs targeting specific microsatellites in the chickpea genome were chosen from literature, sorted out with multiple in-silico tests, and 20 primer pairs were tested in the lab. With one singular primer pair most of the 24 cultivars can be distinguished. However, most of the remaining cultivars can be differentiated with other primer pairs.

For SNPs, primer pairs were newly designed. With the aim to obtain more complex melt curves, we tried to target multiple SNPs with one and the same primer pair. The locations of SNPs being in close vicinity and their surrounding bases were identified with the help of literature. Nine primer pairs passed the *in silico* tests and were tested in the lab. To gain more insight into the melting curves, the amplicons containing SNPs were sequenced. Similar to the assays targeting microsatellites, one primer pair distinguished the majority of the 24 cultivars and most of the remaining samples are complemented by other tested primer pairs.

By combining the results of both types of molecular markers all cultivars except four pairs can be differentiated.

Zusammenfassung

Der Anbau von Kichererbsen wird mit dem fortschreitenden Klimawandel immer interessanter, da Kichererbsen in wärmeren Umgebungen besser wachsen und weniger empfindlich gegenüber Wasserknappheit sind. Zusätzlich sind Kichererbsen eine gute Proteinquelle, enthalten lösliche und unlösliche Ballaststoffe, weisen einen niedrigen Fett- und Natriumgehalt auf und enthalten kein Cholesterin. Mit dem steigenden Trend zu veganen Produkten und Fleischalternativen werden Lebensmittel mit Kichererbsen in europäischen Supermärkten immer beliebter.

Diese Studie zielte darauf ab, 24 Kichererbsensorten, die den drei Typen Desi, Kabuli und Gulabi angehören, zu differenzieren, wobei die meisten aus Europa stammen. Abgesehen von den visuellen Unterschieden zwischen Desi, Kabuli und Gulabi ist keine morphologische Unterscheidung zwischen den Sorten innerhalb eines dieser drei Typen möglich. Wir entwickelten DNA-basierte Assays beruhend auf der Polymerase-Kettenreaktion (PCR), um Kichererbsen-DNA mit Mikrosatelliten oder Einzelnukleotid-Polymorphismen (SNPs) zu amplifizieren, gefolgt von einer hochauflösenden Schmelzkurvenanalyse (HRM-analysis).

Mehrere Primerpaare, die auf spezifische Mikrosatelliten im Kichererbsengenom abzielen, wurden aus der Literatur ausgewählt, durchliefen mehrere *in silico* Tests und 20 Primerpaare wurden im Labor getestet. Mit einem einzelnen Primerpaar können die meisten der 24 Sorten unterschieden werden. Allerdings können die meisten der verbleibenden Sorten mit anderen Primerpaaren differenziert werden.

Für SNPs wurden Primerpaare neu entworfen. Mit dem Ziel, komplexere Schmelzkurven zu erhalten, versuchten wir, mehrere SNPs mit ein und demselben Primerpaar zu erfassen. Nah beieinanderliegende SNPs und deren umliegende Basen wurden mithilfe der Literatur identifiziert. Neun Primerpaare bestanden die *in silico* Tests und wurden im Labor getestet. Um mehr Einblick in die Schmelzkurven zu erhalten, wurden die Amplikons mit SNPs sequenziert. Ähnlich wie bei den Assays, die auf Mikrosatelliten abzielen, unterschied ein Primerpaar die Mehrheit der 24 Sorten und die meisten der verbleibenden Proben werden durch andere getestete Primerpaare differenziert.

Durch die Kombination der Ergebnisse beider Typen von molekularen Markern können alle Sorten außer vier Paare differenziert werden.

Contents

Acknowledgments	I
Abstract	III
Zusammenfassung	V
1 Introduction	1
1.1 Authenticity	2
1.2 Analytics	3
2 Aims	5
3 Theory	6
3.1 <i>In silico</i> primer tests & primer design	6
3.2 DNA extraction	7
3.3 Fluorimetry	8
3.4 Gel electrophoresis	9
3.5 Polymerase chain reaction	10
3.5.1 Principle	10
3.5.2 Main components	11
3.5.3 Real-time PCR (qPCR)	13
3.5.4 Advantages & limitations	14
3.6 High resolution melting analysis	15
3.6.1 Principle	15
3.6.2 Evaluation of melting curves	16
3.6.3 Advantages & limitations	17
3.7 Pyrosequencing (PSQ)	18
3.7.1 Principle	18
3.7.2 Advantages & limitations	19
3.8 Microsatellites	19
3.9 Single nucleotide polymorphism	19
4 Experimental part	21
4.1 Samples	21
4.2 <i>In silico</i> primer tests & primer design	24
4.2.1 <i>In silico</i> primer tests	24
4.2.2 Optimization	31
4.2.3 Primer design	34

Contents

4.3	DNA extraction	35
4.4	Fluorimetry	36
4.5	Gel electrophoresis	37
4.6	Primer dilution & aliquotation	38
4.7	Polymerase chain reaction	39
4.7.1	Optimization	41
4.8	High resolution melting analysis	41
4.9	Pyrosequencing	43
5	Results & discussion	45
5.1	Microsatellites	45
5.1.1	Suitability of primer pairs	45
5.1.2	Chickpea cultivar differentiation based on microsatellites	57
5.2	Single nucleotide polymorphism	66
5.2.1	Suitability of primer pairs	66
5.2.2	Chickpea cultivar differentiation based on SNPs	68
5.2.3	Sequencing results	74
5.3	Combination of results based on microsatellites & SNPs	79
6	Conclusion	80
7	Materials	82
7.0.1	Laboratory equipment	82
7.0.2	Software	82
7.0.3	Solutions & kits	83
7.0.4	Webserver	83
7.0.5	Other utensils	83
	Bibliography	84
	Appendix	90
	Abbreviations	90
	List of Tables	92
	List of Figures	93
	Raw data	100

1 Introduction

Climate change is already in full swing. Crops are exposed to environmental stresses, including longer periods of drought, extreme fluctuations in temperature, increasing carbon dioxide concentrations, heavy rain leading to floods and waterlogging, and changes in pH and metal toxicity [1]. Scientists expect that a global warming of 1.5 degrees Celsius could be reached as early as the 2030s to 2050s [2]. With the increasingly growing world population and rising temperatures, the food supply chain must be maintained. Therefore, it is important to look for ways and means to feed such a large global population. In the coming years, the cultivation of crops that can cope well with high temperatures and dry areas will become increasingly important.

Chickpeas fall into this category. In addition to their excellent resistance to warmer temperatures, they can also cope well with water limitations [3]. Currently, a lot of effort is put into generating even more biotic and abiotic stress-resistant chickpea cultivars [4]. Chickpeas are already gaining popularity due to the growing trend of vegetarian and vegan substitute products [5]. For example, chickpeas are the main ingredient in products such as hummus and falafel and are often used as a meat substitute and rarely as a coffee surrogate [5], [6]. Chickpeas are an excellent source of easily digestible protein, as 80 % of their dry weight is constituted of proteins and carbohydrates [7], [8]. Chickpeas also contain a lot of dietary fibre, are low in fat with a high content of unsaturated fatty acids, have almost no cholesterol and are rich in minerals and vitamins [9], [10].

According to Li *et al.* the general advantages of legumes can be summarized as follows [11]:

- high content of proteins and essential amino acids
- high content of soluble dietary fibres
- increased content of resistant starch after cooking
- protease- and α -amylase inhibitors to slow down the digestion of proteins and carbohydrates, which leads to longer saturation

Chickpeas, scientifically known as *Cicer arietinum*, are an annual legume and belong to the family Fabaceae [6]. Chickpeas are one of the earliest cultivated legumes as they were domesticated over 11,000 years ago and are currently the world's second most important food legume [8], [12]. Although they belong to one of the earliest cultivated plants, chickpeas have not received the same attention in the research field as other crops such as wheat or barley [9]. *Cicer reticulatum* is considered to be the wild progenitor of the

1 Introduction

cultivated chickpea [6], [13]. They are diploid organisms with a total of 16 chromosomes arranged in eight pairs. Almost all known *Cicer* species are self-pollinating [6], [14].

It is suspected that cultivated chickpeas share a narrow genetic base. Currently a lot of effort is put into breeding more profitable cultivars, which are also more resistant to temperature fluctuations or other biotic and abiotic stresses. As the cultivars are closely related among themselves, improvement of yield and tolerance to various external factors are rather slow in chickpeas despite extensive breeding efforts [15], [16].

According to the FAO 18.1 million tonnes of chickpeas were produced worldwide on a cultivated area of 14.8 million hectares in the year 2022 [17]. 70 % of the global production lies in India followed by Australia, Turkey, Russia and Myanmar [17]. The chickpea plant reaches a height of around 30 - 50 cm and one seedpod usually contains between one to three chickpeas. [18], [19].

When it comes to harvested area and production, legumes are the second most important agricultural family right after cereals. They are commonly used as food including processed vegetable oil and animal feed. Additionally, they improve the sustainability of agriculture by enriching the soil with nitrogen through nitrogen fixation with symbiotic bacteria. Thus, less fertilizer is required [20].

There are multiple chickpea types native to different parts of the world. The two main types are called kabuli and desi [6], [14]. The kabuli type is common in the Mediterranean, including southern Europe, as well as northern Africa, South America and India. Chickpeas belonging to the kabuli type are of a lighter color similar to beige, rather large in size in comparison to the other types and have a smooth coat. The desi type is mainly native to India. Chickpeas of the desi type are smaller in size, have a dark brown color and a rough coat [6], [21]. Cultivars belonging to the kabuli type have white flowers and lack anthocyanin pigmentation on their stems, while desi types have pink flowers with high anthocyanin pigmentation on their stems [22]. The appearance of the gulabi type is similar to desi chickpeas, as the seeds are smaller and have a darker color. The gulabi type is a desi subtype, but lesser known than the kabuli and desi types and also not well researched. In comparison to kabuli and desi, cultivars belonging to the gulabi type tend to require more days until they reach physiological maturity [23].

For this thesis, multiple chickpea cultivars belonging to the three types, desi, kabuli and gulabi, were analyzed. The aim was to develop a new authentication method, as a morphological distinction between cultivars belonging to the same type proves to be challenging.

1.1 Authenticity

Food products are deemed authentic when their properties align with the information stated on the label and are free of adulteration. Accurate labelling is crucial for informing consumers about the nature of the products they are buying and consuming [24].

The legal framework established by the EU to combat food fraud is listed under EC No. 178/2002 [25]. It ensures a high level of protection of human life and health by protecting consumer interests including fair practices in food trade. It includes regulations regarding

the labeling and presentation of food including the advertisement. Some commonly adulterated food products include olive oil, honey, milk and dairy products, spices and herbs, seafood, wine, fruit juices, vanilla extract and food advertised as organic despite not meeting organic farming standards. While there are specific guidelines for verifying the authenticity of olive oil, honey, and wine, other food categories are lacking in this regard, which makes the assessment more complicated and challenging [24].

Various sectors of the food industry are impacted by food fraud, which is a significant issue on a global level. Adulteration can involve geographical origin, species falsification, production methods and/or the composition of the product [24]. Especially economically motivated adulteration takes the lead which involves intentional substitution, addition, tampering or mislabelling food, ingredients or packaging for economic profit [26]. Therefore, the development of methods to determine the authenticity of various foodstuffs is required. Especially the analysis of processed products is challenging. Thus, highly sensitive analytical methods are required to verify authenticity.

1.2 Analytics

When it comes to plants, there are multiple methods to confirm authenticity, including observing the metabolite profile. However, this method is not very reliable as metabolites are highly influenced by external factors such as temperature, infections, light and storage conditions [20]. As an alternative, morphological analysis which involves the examination of physical traits of plants (leaf shape, flower, color, growth habits) can provide important information for cultivar differentiation [27]. Morphological analysis can only be considered for non processed food and proves rather challenging for chickpeas. As already mentioned before, morphological distinction is difficult between chickpea cultivars belonging to the same type. Therefore, DNA-based methods are preferred.

There are various ways in which DNA can be used to determine authenticity. First and foremost, DNA fingerprinting has become a well-established technique in the last decades. It involves analyzing a unique DNA sequence of a plant to create a genetic fingerprint. Different markers such as microsatellites, also known as simple sequence repeats (SSR), and single nucleotide polymorphism (SNP) are commonly used [28], [29].

Microsatellites are short, highly repetitive DNA sequences made up of one to six nucleotides that can repeat 10 - 1,000 times and usually appear at the same loci [28], [30], [31]. Their length varies significantly among different cultivars of the same species, exhibiting high allelic variation due to slipped-strand mispairing, gene deletion or duplication events, and other less understood mechanisms [30]–[32].

SNP refers to the variation of a single nucleotide at a specific locus in the genome [33]. SNPs can be inherited and are not evenly distributed throughout the genome, with certain areas known as SNP hotspots. They can occur in both coding and non-coding regions. Most SNPs involve the substitution of cytosine (C) with thymine (T) or guanine (G) with adenine (A) [33], [34]. SNPs are also considered successful point mutations, as their presence has been established in a certain percentage of the population [33], [35].

1 Introduction

In order to use these markers for cultivar differentiation, they are first amplified with polymerase chain reaction (PCR) using specific primers targeting the flanking regions [32].

Following the PCR, a high resolution melting (HRM) analysis can be conducted. HRM analysis is a method used for cultivar differentiation, as it provides a sensitive, precise and non-destructive method for detecting genetic variations in DNA sequences. HRM works by monitoring the melting behaviour of double-stranded DNA fragments as they are heated. Different cultivars often have distinct genetic variations, such as SNPs which cause differences in their melting profiles [36].

As an alternative approach, next-generation sequencing (NGS) methods may provide comprehensive and detailed genomic information. Entire genomes or specific regions can be sequenced in high resolution to detect small genetic variations such as insertions, deletions and SNPs. By comparing the results with databases, NGS facilitates the identification of plant varieties with high precision to ensure authenticity. Additionally, extensive sequencing efforts for chickpeas are undertaken to understand the genetic diversity and improve traits like disease resistance and yield in chickpea breeding programs [37].

While HRM was used prior to differentiate various legume types from each other, including soybean, lentils and chickpeas, no HRM-based method has been established so far targeting different chickpea cultivars [20]. This thesis focuses on the differentiation of various chickpea cultivars using PCR to amplify regions containing microsatellites or multiple SNPs as markers followed by HRM analysis.

2 Aims

Food fraud remains a significant global issue affecting various sectors of the food industry. Consequently, developing methods to authenticate various foodstuffs is crucial.

This thesis is a cooperation between the University of Vienna and the Austrian Agency for Health and Food Safety (AGES) and aims to develop multiple DNA-based assays involving PCR followed by HRM analysis to differentiate 24 different chickpea cultivars. HRM is a technique based on the different melting behaviour of double-stranded DNA fragments. As target sequence different microsatellites and SNPs present in the chickpea genome were chosen.

Chickpeas are becoming more popular due to their increased use in vegetarian and vegan products. Additionally, chickpeas might offer a stable food source regarding the climate change due to their resistance to warmer temperatures and drought.

While a HRM-based method distinguishing different legume types was already established, there is currently no HRM-based method to differentiate different chickpea cultivars. Different primer pairs targeting microsatellites taken from the literature are tested for their suitability to differentiate 24 different chickpea cultivars. A different set of primer pairs, which are newly designed and targeting multiple SNPs, are also tested for their differentiation properties. Amplicons containing SNPs are subsequently sequenced using pyrosequencing to gain more insight into the melting behaviour of the different cultivars.

3 Theory

3.1 *In silico* primer tests & primer design

Primers are short single-stranded DNA sequences which serve as an anchor point for the DNA polymerase during PCR. They set the target region, which is to be amplified [38], [39]. Primer design is a crucial step in which a lot of time should be invested, as well-designed primers enhance the quality of the PCR result [40]. Poorly chosen primers might lead to little or no PCR product or bind to unspecific regions [39]. Typically, DNA sequences, which are characteristic for the respective organism, are chosen as the target region for cultivar differentiation [41]. For efficient amplification primer pairs are used, which consist of a forward and a reverse primer. The forward primer binds to the upper and the reverse to the lower strand [39].

In order to optimize the primers for a specific target region, *in silico* tests can be carried out. *In silico* tests are computational tools which predict the probability of dimer formation, secondary structures or the melting temperature of primer pairs and give more insight into how they might behave *in vitro* (view section 4.2.1).

Primer pairs should have approximately the same length. 18 - 24 nucleotides are a good size, as too short primers might not bind specifically and longer primers might form secondary structures, which complicate the binding to the template DNA [39]. Primer length also directly influences the annealing temperature, which ideally lies between 52 - 65 °C [39], [40]. Especially the 3' end of the primer has to be specific and should not contain a lot of A and T consecutively, as they only form two hydrogen bonds and are thus not very stable [39]. Additionally, the primer pair sequences should not be complementary to each other as the formation of primer dimers should be avoided. Primer dimers consume important resources for the PCR such as free deoxynucleotides and polymerase activity, which results in a competition for PCR reagents [39]. It is also important that the primer binding site does not contain any frequently occurring genetic variants such as SNPs, as this negatively impacts the binding of the primer to the template. Ideally, the content of G and C in the primer should fall between 40 and 60 % [39]. A higher GC content greatly increases the melting and annealing temperature and thus the temperature needed to remove the primer from the template, whereas a too low GC content in the primer leads to rather unstable binding. The melting temperature between forward and reverse primer should only vary by maximum 1 °C, as otherwise the discrepancy of the optimal binding conditions is too high.

The resulting amplicon length for microsatellites should reach a maximum length of 250 - 300 bp, which is still usable for HRM analysis [28]. For SNPs, however, the shorter the amplicon, the better, as in longer amplicons other motifs apart from the present SNPs might impact the melting behaviour [36], [40].

3.2 DNA extraction

DNA extraction is a process used to isolate DNA from cells or tissues. The protocol involves multiple steps starting with sample homogenization, cell lysis, protein denaturation and precipitation, separation of DNA, DNA purification, washing steps and resuspension [41], [42].

Homogenization involves the crushing, mashing and pulverizing of the whole sample to generate a homogenous powder or pulp [41], [42]. This step might involve the use of a mortar and pestle, liquid nitrogen, presses or other devices [42]. After successful homogenization, the sample material is accessible for the respective analysis method.

There are multiple methods to obtain intact DNA from plant source materials.

Phenol extraction involves the addition of phenol or a phenol-chloroform mixture after cell lysis. Additionally, protease to degrade present proteins can be added. After centrifugation, the formation of multiple phases can be observed. At the bottom of the vial is the organic phase containing lipids, followed by a slim phase of denatured proteins. The aqueous phase containing the DNA is located at the top [43].

The CTAB method involves the use of a cetyltrimethylammonium bromide (CTAB) buffer, which is added to the cell lysate (Figure 1). The positively charged CTAB forms a complex with DNA, which is not soluble and precipitates into a pellet after centrifugation. The supernatant containing carbohydrates and proteins is discarded. The CTAB-DNA complex can be brought back in solution by adding a high concentrated NaCl solution. After an additional precipitation step involving alcohol (e.g. isopropanol), the pure DNA is obtained [41].

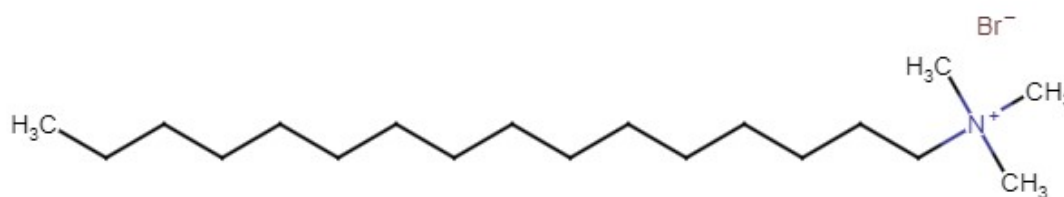


Figure 1: Chemical structure of CTAB drawn using Chemical Sketch Tool.

3 Theory

The main method used in this master thesis involves the use of commercial silica columns. First, DNA is released using sodium dodecyl sulfate (SDS) (Figure 2). SDS is a negatively charged detergent, which unfolds proteins by binding to their hydrophobic regions. Proteins are removed from their original association and are rendered soluble [38]. SDS disrupts the cell membrane and nuclear envelope leading to cell lysis and thus releasing the DNA.

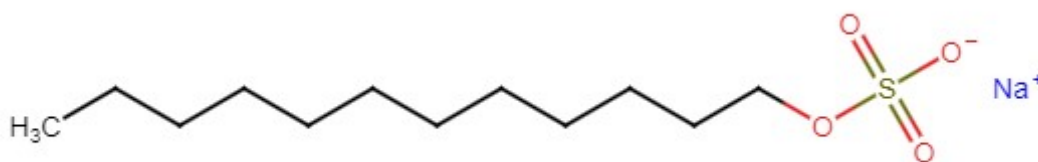


Figure 2: Chemical structure of SDS drawn using Chemical Sketch Tool.

The lysate is then filtered using a special column to clarify the solution and remove polysaccharides, cellular debris, and other contaminants. A second type of column is used to bind DNA after increasing the salt concentration of the filtrate. Only under high salt concentrations can DNA be bound by forming salt bridges with the free hydroxy groups of the silica column. The column is washed multiple times with different washing solutions with high salt concentrations. DNA is eluted by reducing the salt concentration and thus destroying the salt bridges [42].

3.3 Fluorimetry

Fluorimetry is a sensitive method for the quantitative analysis of substances based on the principle of fluorescence. A light source emits light of a certain wavelength, which is absorbed by molecules in the sample. During this process electrons are lifted on a higher energetic level. Shortly after, the electrons fall back to their ground level, releasing energy in the form of light. The emitted wavelength is longer and has a lower energy level than the original wavelength from the light source [41].

In order to measure DNA with fluorimetry fluorescent dyes are required, as DNA itself is not fluorescent. The fluorescent dye binds specifically to double-stranded DNA and is only fluorescent in its bound state. Thus, the fluorescence signal increases proportionally with the DNA concentration. To calculate the DNA concentration an external calibration after linear regression is used [41], [44].

3.4 Gel electrophoresis

Charged molecules such as DNA can be separated in an electric field. This is the base principle of electrophoretic methods. Due to differences in charge and size the molecules wander in the gel at different paces, which results in their separation. For DNA gel electrophoresis, agarose gels are used, as they are chemically inert, and their pore size and structure can be influenced by the amount of added agarose [41].

Gels are placed in an electrophoresis chamber with the cathode on top and the anode at the bottom. The chamber is filled with a buffer solution until the gel is completely covered to ensure a connection between the gel and the electrodes. Samples have to be mixed with glycerin to add some weight to the samples and make loading of the gel pockets easier. Usually, the glycerin mixture also contains a colorant to make it easier to observe the progression of the DNA in the gel during electrophoresis [41].

DNA is negatively charged due to its phosphate backbone. When exposed to an electric field, the DNA thus migrates to the positively charged anode. Depending on the length of the DNA fragments they move at different paces. Longer DNA fragments are slower and take more time to diffuse through the gel, while smaller fragments can easily traverse the pores of the gel. To estimate the size of the separated fragments commercially available DNA markers are used, which contain a mixture of DNA fragments with specific lengths [41].

In order to make the DNA fragments visible, an intercalating fluorescent dye is necessary. Ethidium bromide was a frequently used reagent to stain DNA fragments, but due to its mutagenic properties, the usage of different color reagents is recommended [45]. After staining the gel, the DNA fragments are visible under UV light [41].

Gel electrophoresis can give more insight into a PCR result. Since the approximate length of the amplicon is known, the gel shows if the primers were specifically bound and only the target region was amplified. If multiple DNA fragments of highly variant sizes are visible on the gel, then different genomic regions were amplified, or if very short fragments are present, primer dimers were formed (Figure 3).

3 Theory

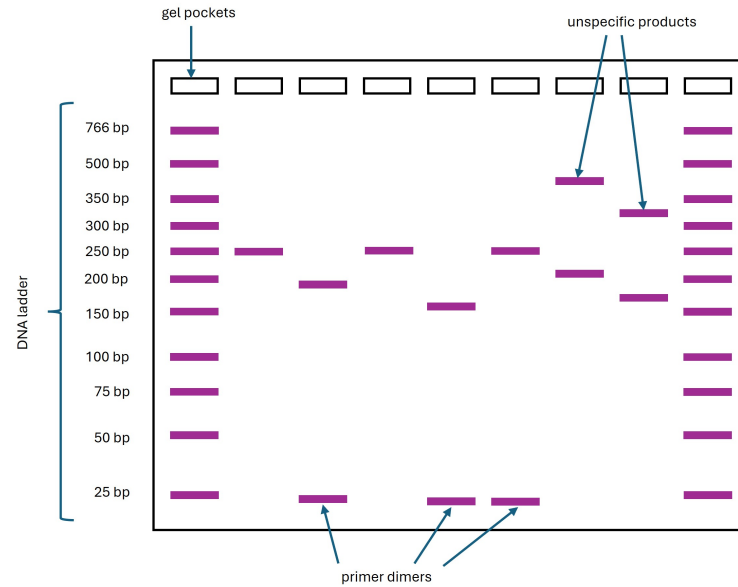


Figure 3: Simplified schematic of the result of an agarose gel electrophoresis. Purple lines represent DNA of various sizes. First and last pocket (lane) of the gel are used for the DNA ladder. The other lanes are filled with sample DNA. Unspecific products and primer dimers are marked with arrows.

3.5 Polymerase chain reaction

PCR is a method to amplify specific DNA sequences. In just a few hours millions of copies can be generated. This method was developed in 1983 by Kary Mullis and has been of high importance ever since [38].

3.5.1 Principle

PCR involves multiple steps, which are shown in Figure 4. First, DNA which is to be amplified has to be denatured. Therefore a high temperature (95 °C) is applied. Additionally, during this step, the DNA polymerase is activated [38], [41].

As the DNA is now single-stranded, the temperature is reduced to the optimal annealing temperature of the primer. This temperature is highly primer specific and is prior tested *in silico* (view section 3.1) as well as optimized *in vitro* (view section 4.2.2).

Once the primers are bound to their specific target regions, the temperature is increased to 72 °C, which is the optimal working temperature for the Taq polymerase. The polymerase elongates the primers by adding free nucleotides to their 3'-end. Upon completion of this step, this whole cycle is repeated up to 50 times, theoretically resulting in the exponential generation of millions of DNA copies [38], [41].

The last elongation step is elongated to 5 - 15 minutes, which ensures that partially synthesized DNA fragments are fully completed by the DNA polymerase. This limits the formation of unspecific products [46].

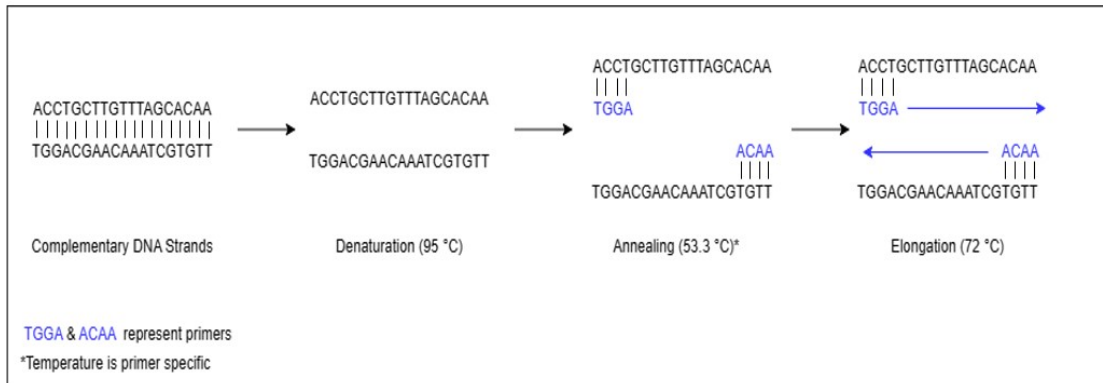


Figure 4: Simplified representation of one PCR cycle.

PCR can be carried out as a singleplex or multiplex assay. Singleplex assays work with one primer pair and a fluorescent probe targeting a single template. Multiplex assays aim at two or more templates using two or more primer pairs and a corresponding number of fluorescent probes in a single reaction approach. Multiplex assays are more time and cost-efficient, use less sample material and lower the risk of potential contamination. However, it is more difficult to establish a functioning multiplex method. All used primer pairs and probes need to have the same annealing temperature, should not form any dimers or secondary structures and the templates must be amplified in the same ratio [47].

3.5.2 Main components

For PCR, a so-called master mix is necessary, which contains all the main components needed for a successful reaction. In this section, every component is described in detail:

Buffer

The buffer solution provides a stable environment for the reaction and keeps the pH levels steady. It generates a suitable environment for the DNA polymerase and supports its proper functionality [41]. The composition of the buffer can vary. There are multiple commercially available buffer or master mix solutions whose exact composition is a company secret. Usually, the further listed components (except for primer and template) are already included in commercially available master mix solutions.

Magnesium chloride

Magnesium chloride is a crucial co-factor for the DNA polymerase as it is necessary for the polymerase's catalytic activity. In some cases, an increase in magnesium chloride concentration can provide better results [41].

Free deoxynucleotides

Deoxynucleotides or also called deoxynucleotide triphosphates (dNTPs) are the building blocks needed to synthesize new DNA strands during amplification. There are four types of dNTPs, which are used in PCR corresponding to one of the four nucleotide bases of the DNA [38], [41]:

- deoxyadenosine triphosphate (dATP)
- deoxythymidine triphosphate (dTTP)
- deoxycytidine triphosphate (dCTP)
- deoxyguanosine triphosphate (dGTP)

DNA polymerase

DNA polymerase is an enzyme which is not only responsible for DNA replication *in vivo* but can also be used for PCR to generate millions of copies of a certain target DNA *in vitro*. It synthesizes new DNA strands by adding nucleotides to the 3'-end of an annealed primer. For PCR the most commonly used polymerase is the Taq polymerase. Its name is derived from the organism *Thermus aquaticus*, a bacterium which thrives in hot springs, from which the polymerase is isolated [39]. The Taq polymerase is especially suitable for PCR as it is able to withstand the high temperatures required for DNA denaturation under which standard proteins would denature [38].

Fluorescent dye

For real-time monitoring of the reaction, fluorescent dyes can be added to the master mix. Commonly used dyes are SYBR[®] Green or EvaGreen[®]. They bind to double-stranded DNA and only, when bound, emit a fluorescence signal. The increase in fluorescence is directly proportional to the number of generated DNA copies. However, fluorescent dyes cannot discriminate between the double-stranded template or primer dimers, which might lead to a decrease in specificity [41], [47], [48]. Compared to SYBR[®] green, which is a potentially mutagenic cyanine dye, EvaGreen[®] is less toxic and non-mutagenic, making it safer to handle, and it is considered to be more environmentally friendly [48]–[50]. SYBR[®] green has limited stability, and if present in higher concentration, it can inhibit the polymerase. Additionally, EvaGreen[®] is more suitable for HRM analysis, as it shows less inhibitory effects [48].

As an alternative to fluorescent dyes fluorescent marked probes such as the TaqMan[™] probe can be used [47], [51]. Probes are specific for the target DNA sequence and consist of a short complementary DNA fragment with a linked fluorescent dye and a quencher. Once the probe is bound to its template, the polymerase removes the building blocks of the probe with its 5'-3' exonuclease activity, thus releasing the fluorescent dye. As the dye is no longer in close proximity to its quencher an increase in fluorescence can be detected [41], [51].

Primer

Primers are short single-stranded DNA sequences. Primers are not included in commercially available master mix solutions. They are designed to target specific templates of interest. For further information view section [3.1](#).

Template

Template is the target DNA sequence which is to be amplified during PCR. For successful amplification the purity, quality and quantity of the DNA template are crucial. Similar to the primer, the template is not included in commercially available master mix solutions.

Thermocycler

A thermocycler is the device used to carry out a PCR. It is able to cycle between different temperatures at predefined time intervals in order to facilitate the amplification of specific DNA fragments [\[41\]](#). Multiple companies offer thermocyclers. They all consist of a thermal block with precise temperature control to rapidly heat or cool samples and a programmable interface to control the cycling parameters.

3.5.3 Real-time PCR (qPCR)

Real-time PCR or also called quantitative PCR (qPCR) is a variation of the traditional PCR method which allows the real-time monitoring of the DNA amplification process. The increase in DNA concentration is measured by an increase in fluorescence signal due to the usage of fluorescent dyes, which bind to double-stranded DNA or fluorescent probes complementary to the target sequence. This results in the formation of an amplification curve (Figure [5](#)), where the increasing fluorescence signal is plotted against the number of cycles [\[41\]](#), [\[51\]](#).

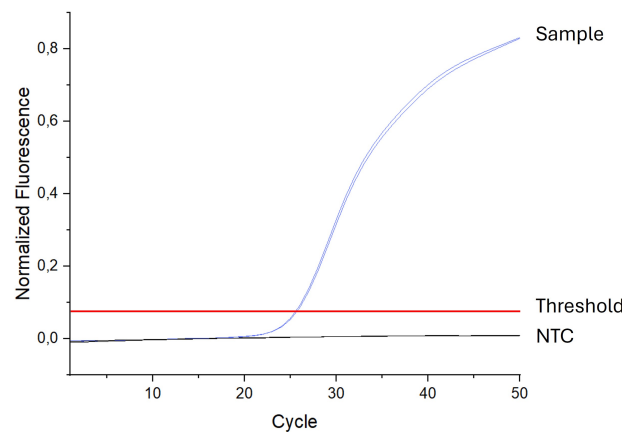


Figure 5: Simplified schematic of an amplification curve generated from a DNA sample, as well as a no template control (NTC). The threshold is marked as a red line. Graphic modified from Stellberger *et al.* [\[52\]](#).

The cycle number at which the created fluorescence signal reaches a detectable threshold level exceeding the background noise is called the cycle threshold (CT) value [51]. It is an important parameter when it comes to the quantification of the target DNA. Usually, for qPCR, the CT-value is determined by the thermocycler software, but the threshold can also be adjusted manually. The CT-value also represents the cycle number at which the DNA amplification reaches exponential growth. It is also inversely related to the target DNA concentration in the sample, as a higher DNA concentration at the beginning of the PCR leads to a lower CT-value and *vice versa*. By comparing the CT-values of unknown samples to a standard curve generated from known concentrations of the target DNA, the quantitative amount of the target DNA in the unknown samples can be determined. It also allows for comparison of samples across different experiments and laboratories [41].

3.5.4 Advantages & limitations

PCR offers many advantages. It provides nearly exponential amplification of the DNA template and generates millions of copies in just a few hours, which makes it very time-efficient. PCR is a highly sensitive method with a very low limit of detection, as already a small number of templates is sufficient to generate millions of copies. When the primers are properly designed PCR also provides a very high specificity [38], [41].

Just as every method PCR also suffers from limitations. The amplification of PCR is only exponential in theory, as it is limited by the amount of free deoxynucleotides present in the master mix as well as by the increasing concentration of pyrophosphate, which is cleaved from each nucleotide upon its insertion into a newly generated DNA strand. Pyrophosphate is a known polymerase inhibitor, and its increase during progressing PCR cycles negatively impacts the activity of the polymerase. Thus, the number of PCR cycles which can be carried out efficiently is limited [41].

Apart from pyrophosphate there are also other polymerase inhibitors which stem from the sample matrix. Commonly known polymerase inhibitors present in plants include polysaccharides and polyphenols. This issue can be overcome with proper sample preparation and DNA extraction steps to separate the DNA from potential inhibitory components. Since PCR is a rather sensitive method, DNA samples can be diluted with DNase-free water to further decrease the concentration of the sample matrix [41].

With progressing cycle numbers, the polymerase also runs through multiple heating and cooling steps. Even though the Taq polymerase which is more sturdy towards higher temperatures is used, the denaturation temperature of 95 °C still lies above its temperature optimum of 72 °C which might lead to a decrease in polymerase activity during later cycles and increases the probability of the formation of unwanted by-products [41].

Additionally, PCR is rather susceptible towards contamination. Samples can be cross-contaminated, which might already happen during DNA extraction. The contamination risk after successful amplification is even greater, as tiny droplets of the reaction mixture already contain hundreds of thousands of copies of the amplified DNA. Even contamination through aerosols is possible. Thus, careful operation and proper work etiquette are necessary, and workspace separation is recommended [41].

3.6 High resolution melting analysis

HRM analysis is a method used to detect different genetic variations with high precision and sensitivity. It is suitable to detect differences in amplicon length as well as mutations and polymorphisms [36], [40], [53].

HRM is performed right after the completion of all PCR cycles. Not even a different device is needed, as most thermocyclers are capable of performing a HRM [36], [53].

3.6.1 Principle

During HRM the temperature is gradually increased in small steps (0.1 °C every two seconds) starting from usually 60 - 65 °C up to 90 - 95 °C. This leads to the gradual denaturation of the amplified DNA strands. As the double-stranded DNA is bound by a fluorescent dye like EvaGreen® emitting a high fluorescence signal, the denaturation of said sample DNA also leads to a continuous decrease in fluorescence, as EvaGreen® can no longer bind. This leads to the formation of melting curves as shown in Figure 6. At the turning point of the curve, which also marks the melting temperature, half of the present sample DNA is melted, while the other half is still double-stranded and bound to EvaGreen®. At the end of the HRM, all sample DNA is denatured [36], [53].

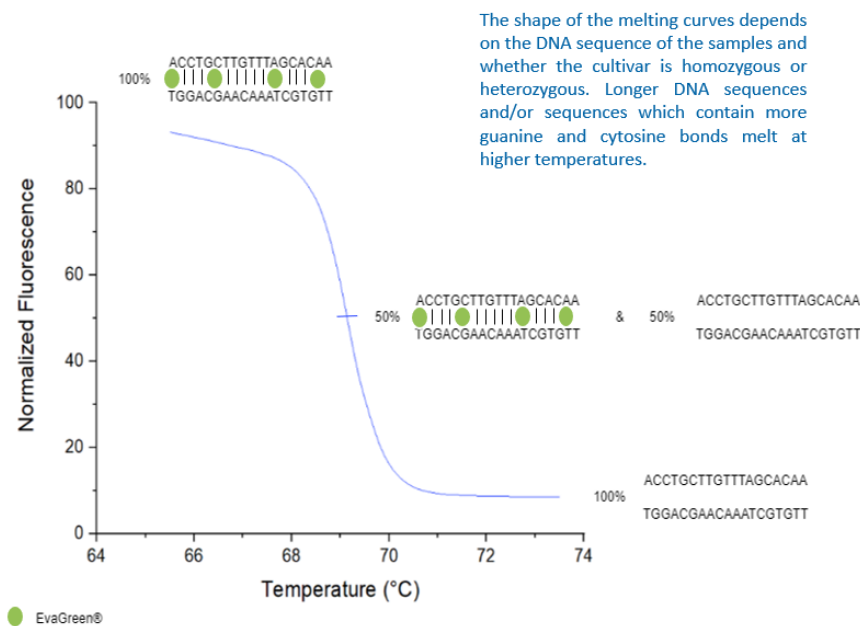


Figure 6: This figure shows a standard melting curve generated by plotting the normalized fluorescence signal against the temperature. While the DNA is double-stranded, a high fluorescence signal is detected due to the bound fluorescent dye EvaGreen®. By continuously increasing the temperature the DNA denatures, thus leading to a decrease in fluorescence signal as EvaGreen® can no longer bind to single-stranded DNA.

3 Theory

DNA strands show different melting behaviours based on their length and on their sequence [36]. Longer DNA sequences denature at higher temperatures compared to shorter sequences. Additionally, DNA sequences which have a higher content in G and C will melt at higher temperatures as they form three hydrogen bonds and are thus more stable compared to A and T rich sequences, which only form two hydrogen bonds [36]. This can be used to visualize polymorphisms in the sample DNA, as already the replacement of e.g. T with G is enough to lead to a higher melting temperature.

HRM is a suitable tool to check whether the desired product was amplified during PCR and to differentiate closely related species by detecting small differences in their DNA sequence.

3.6.2 Evaluation of melting curves

Melting curves provide a lot of information about the DNA sample. As stated before, the shape of the melting curves depends on the DNA sequence, the amplicon length as well as whether the cultivar is homo- or heterozygous [36]. Also, the formation of primer dimers can be observed during HRM [39]. Since primer dimers are very short sequences compared to the target DNA, they will melt at lower temperatures. However, with a good pre-selection of primers as listed in section 3.1 the chance of primer formation during PCR should be minimized.

There are different ways to display melting curves. By plotting the fluorescence signal against the temperature, the standard melting curve, as shown in Figure 6 is obtained [36]. As an alternative the negative derivative of the fluorescence signal (dF/dT) can be plotted against the temperature, which leads to the formation of melt peaks as shown in Figure 7 [36]. In this depiction, the melting temperature and thus the point where half of the DNA is double-stranded and the other half is single-stranded is marked by the peak maximum.

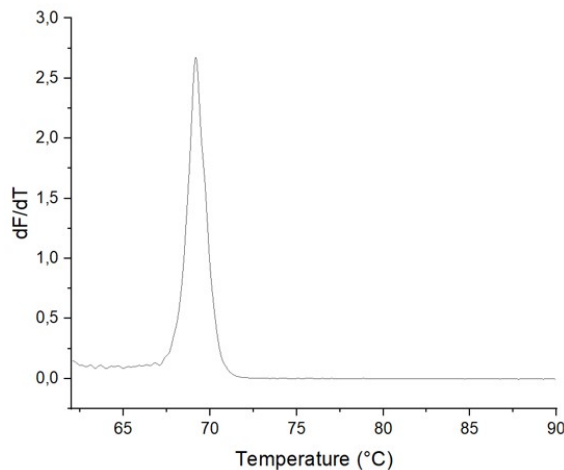


Figure 7: This figure shows an alternative depiction of melting curves in the form of melt peaks. Melt peaks are formed by plotting the negative derivative of the fluorescence signal against the temperature. The peak maximum corresponds to the turning point of the curve.

3.6 High resolution melting analysis

To establish an accurate and reliable interpretation of the data the normalization of melting curves is recommended [36]. Normalization accounts for possible differences in baseline fluorescence levels of the DNA samples due to different initial DNA concentrations, differences in dye binding affinity or instrumental sensitivity. Additionally, melting curves of the different samples are aligned placing them on a uniform scale, which allows for a better comparison between them. It also helps to remove background noise and other potential artefacts, which might affect the accuracy of the result and thus enhance the resolution [36].

The distinct melting curve profile also provides information about whether the analyzed cultivar is homozygous or heterozygous [36], [54]. In homozygous cultivars, both alleles are identical, and thus, they usually produce smooth, symmetrical melting curves and sharp melt peaks since both DNA strands melt at the same temperature. Heterozygous cultivars have two different alleles and thus display more complex melting behaviour, as both alleles do not melt identically. Melting curves tend to be asymmetrical and melt peaks show protruding features [36]. Additionally, heterozygotes can form heteroduplexes due to mismatched base pairs during melting and re-annealing steps, which enhance the complexity of the melting curves [54]. The differences between homozygous and heterozygous samples are displayed in Figure 8.

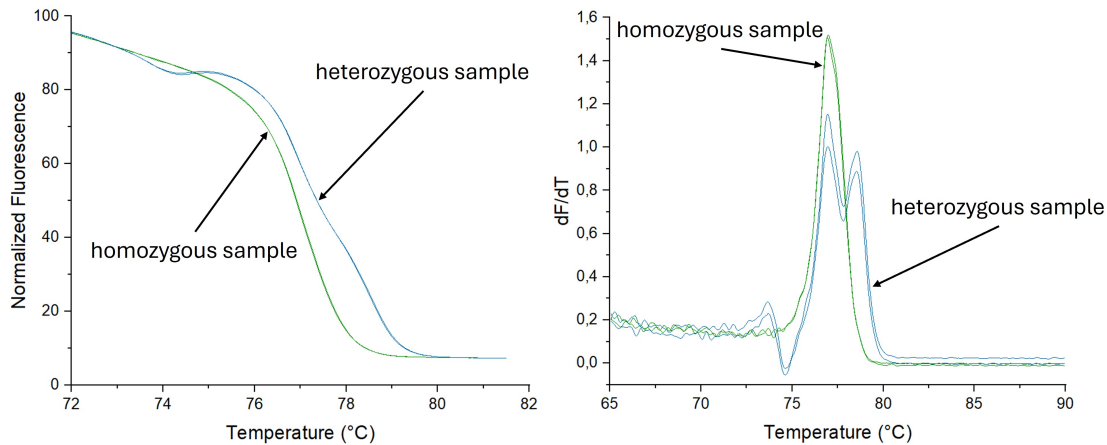


Figure 8: Melting curves and melt peaks of two different chickpea cultivars amplified with primer pair SNP2. Samples were analyzed in duplicates. The sample Rondo (green) is homozygous and shows a symmetrical melting curve as well as one symmetrical melt peak. The sample Pasciá (blue) is heterozygous with an asymmetrical melting curve and a split melt peak.

3.6.3 Advantages & limitations

HRM is a highly sensitive method with high throughput and great specificity as it is possible to detect the deletion, insertion or substitution of one single base pair. Compared to other methods it is fast and cost-effective since it does not require the use of specific DNA probes as it works with fluorescent dyes [40], [53], [54]. Additionally, the method is non-destructive and thus the DNA samples can be used for further analysis [36].

Just as every method, HRM also has some limitations. To obtain reliable results, high purity of the DNA has to be established, as impurities or the presence of matrix components might affect the melting behaviour of the sample DNA [40]. Additionally, in order to analyze the HRM results a specialized software is required.

3.7 Pyrosequencing (PSQ)

Pyrosequencing is a DNA sequencing method which detects the release of pyrophosphate upon the insertion of nucleotides [55]. It is used for genotyping and mutation detection [56].

3.7.1 Principle

Pyrosequencing is based on the "sequencing by synthesis" principle involving the addition of nucleotides to a sequencing primer, which is hybridized to single-stranded PCR-amplified DNA templates. The sequencing process is monitored in real-time, and the incorporation of nucleotides is detected with chemical reactions [56], [57].

One of the primers used for amplification of the target region in the prior PCR is labeled with biotin. During pyrosequencing, the single-stranded template, which is linked to biotin, binds to beads coated with streptavidin, forming a stable complex. Unlabeled strands are removed [56].

Nucleotides are added to the reaction mix one at a time. When the DNA polymerase adds fitting nucleotides to the 3'-end of the primer, a certain concentration of pyrophosphate (PPi) is released corresponding to the number of inserted nucleotides. The enzyme ATP sulfurylase together with adenosine 5' phosphosulfate (APS) converts pyrophosphate to ATP, which is the driving power of another chemical reaction. The enzyme luciferase uses the generated ATP to convert luciferin to oxyluciferin, which leads to the emission of visible light. The intensity of the generated light corresponds directly to the inserted number of nucleotides by the DNA polymerase [55], [57]. With a photomultiplier tube the emitted light is detected and generates a peak in the resulting Pyrogram[®]. Excess nucleotides, which were not incorporated in the DNA and remaining ATP are degraded by the enzyme apyrase before the addition of the next nucleotide. The chemical reactions are summarized in Figure 9. Once the complementary sequence is completed, the nucleotide sequence can be determined by the formed peak pattern of the Pyrogram[®] [57].

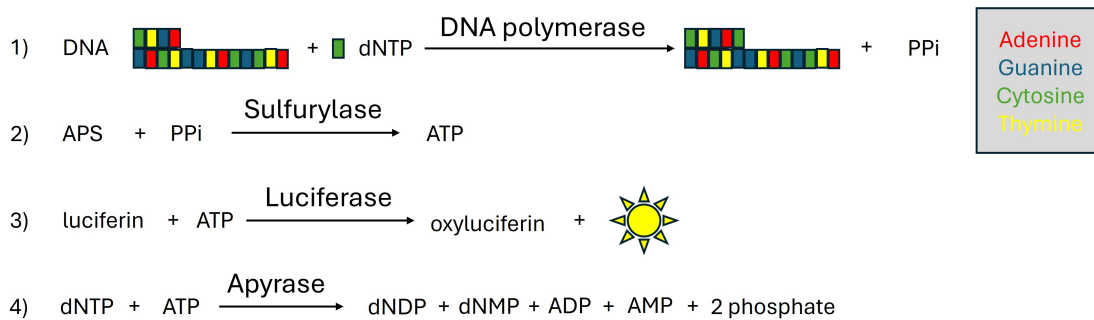


Figure 9: Simplified depiction of the chemical reactions involved during pyrosequencing. Graphic modified from [57].

3.7.2 Advantages & limitations

Pyrosequencing shows high accuracy and high precision in detecting the base sequence of short DNA sequences and specific mutations such as SNPs [55], [56]. It provides real-time analysis of the sequencing data and can be also used to quantify specific DNA sequences, which is useful for allele quantification [55].

Pyrosequencing is mainly limited by the amplicon length, as only 100 - 300 base pairs can be efficiently sequenced. For sequencing larger genomic regions other sequencing methods have to be used [41], [55]. Additionally, the synthesis of homopolymeric regions is problematic due to the nonlinear light response upon the insertion of 5 - 6 identical nucleotides in a row [55].

3.8 Microsatellites

Microsatellites, also known as simple sequence repeats (SSR) or short tandem repeats (STR) are short highly repetitive DNA sequences consisting of one to six nucleotides. These short sequences can repeat themselves 10 - 1000 times and usually appear in the same loci [28], [30], [31].

Microsatellites serve as important markers for cultivar differentiation, as their length is highly variable between different cultivars of the same species with high allelic variation due to strand mispairing, gene deletion or duplication events or other less understood mechanisms [30], [31].

3.9 Single nucleotide polymorphism

SNP describes the variation of a single nucleotide at a specific locus in the genome [33]. They can be inherited. SNPs are not evenly distributed among the genome, but there are SNP hotspots. They might be located in coding or non-coding regions. Most SNPs consist of the exchange of C with T or G with A [33], [34].

3 Theory

SNPs are also considered to be successful point mutations, as their appearance has been established in a certain percentage of the population [33], [35]. There are three possible outcomes for e.g. a C/T SNP regarding diploid organisms:

- C/C: both alleles have a C at the SNP locus
- T/T: both alleles have a T at the SNP locus
- C/T: one allele has C and the other T at the SNP locus

Similar to microsatellites SNPs can be used as markers for cultivar differentiation [33], [35].

4 Experimental part

4.1 Samples

24 different chickpea cultivars from different origins were provided by the Austrian Agency for Health and Food Safety (AGES) in the form of whole chickpeas and chickpea powder. Chickpeas were planted in four sets in a randomized order in Fuchsenbigl in the region of the Marchfeld in Lower Austria in the first week of May 2023. The cultivars are self-pollinating and therefore, cross-breeding during cultivation can be excluded. All samples were harvested together in August 2023. Varietal purity of the samples is assumed.

For this thesis, only the already grounded chickpea powder was used. The chickpeas were ground in two grades by Nathalie Runge. The first grade was rather rough with a particle size of 4 mm, and the second grade produced a fine powder with a particle size of 0.5 mm. Samples were labeled from numbers 1 to 25 including the abbreviation CPS (chickpea sample). Samples CPS3 and CPS4 belonged to the same species but differed in their seeding rate and were analyzed independently.

The provided samples included three different chickpea types: kabuli, desi and gulabi, a desi-subtype. A complete list of all samples sorted by type, including their cultivar name, source and land of origin, can be found in Figure 10. The samples are color-coded, where kabuli types are assigned colors of the blue-green spectrum, desi types are colored red and gulabi purple. Later shown melting curves in the result section follow the color scheme of the samples listed here.

Visually, little to no distinction between chickpeas belonging to the same type is possible. Chickpeas of the kabuli type have beige seeds of a bigger size with a smooth coat (Figure 11). Kabuli and desi types are of a darker color ranging from dark brown to black, have smaller seeds and look rather wrinkled (Figures 12, 13). An exception is the cultivar Olga, which belongs to the gulabi type but has smooth seeds. The ground chickpea powder of the kabuli types was a light yellow, homogenous powder, whereas desi and gulabi had additional dark brown and black specs among the yellow powder.

4 Experimental part

Sample number	Cultivars	Type	Source	Land of Origin
CPS1	Cicerone	Kabuli	Tec2Trade	Italy
CPS2	Vulcano	Kabuli	MyLocalFarm GmbH	Italy
CPS3-CPS4	Flamenco	Kabuli	Strube	France
CPS6	Rondo	Kabuli	Strube	France
CPS7	Donia	Kabuli	Hungarian University of Agriculture and Life Sciences Kaposvár Campus	Hungary
CPS9	Maragiá	Kabuli	Tec2Trade	Italy
CPS10	Analisto	Kabuli	LIDEA	France
CPS11	Jafar	Kabuli	Società Produttori Sementi Spa	Italy
CPS12	Gavdos	Kabuli	Saatbau Linz	Greece
CPS14	Sultano	Kabuli	MyLocalFarm GmbH	Italy
CPS15	Reale	Kabuli	Società Produttori Sementi Spa	Italy
CPS16	Amorgos	Kabuli	Saatbau Linz	Greece
CPS17	Thiva	Kabuli	Saatbau Linz	Greece
CPS18	CDC Orion	Kabuli	LIDEA	USA
CPS19	Elixir	Kabuli	Tec2Trade	Portugal
CPS20	Sokol	Kabuli	BSV Saaten	Russia
CPS22	Badil	Kabuli	Semo Bio GmbH	Spain
CPS23	Pasciá	Kabuli	Tec2Trade	Italy
CPS25	Twist	Kabuli	RWA	France
CPS5	Elmo	Desi	LIDEA	Portugal
CPS13	Bori	Desi	Szarvasi Medicago Kft.	Hungary
CPS24	Katalin	Desi	Szarvasi Medicago Kft.	Hungary
CPS8	Dora	Gulabi	Hungarian University of Agriculture and Life Sciences Kaposvár Campus	Hungary
CPS21	Olga	Gulabi	Research Institute for Fodder Crops Ltd. Troubsko	Czech Republic

Figure 10: List of all cultivar names sorted by chickpea type including their source and land of origin. For experiments sample numbers CPS1 - CPS25 were used. CPS stands for chickpea sample.

4.1 Samples



Figure 11: Picture of the samples Analisto, Elixir and Sokol (from left to right) representing the kabuli type.



Figure 12: Picture of the samples Elmo, Bori and Katalin (from left to right) representing the desi type.



Figure 13: Picture of the samples Dora and Olga (from left to right) representing the gulabi type.

4.2 *In silico* primer tests & primer design

The primer pair sequences targeting microsatellites were chosen from two different papers [28], [30]. All listed primer sequences of both papers were tested *in silico*, however, only primer pairs which met the prior listed criteria in section 3.1 were tested in the lab.

Primer pairs for SNPs were newly designed. To generate more complex melting curves, it was ensured that each amplicon contained at least three SNPs. Table 3 from the paper published by Basu *et al.* was used to find SNPs which are closely spaced next to each other on the chickpea genome [14]. Additionally, the exact location of the SNPs in the genome was listed. This was of great use when looking for the DNA sequence surrounding the SNPs. Table S16 from the paper published by Parween *et al.* included 100 bases upstream and downstream of SNPs found in the chickpea genome, which was used to locate suitable primer sequences in the flanking regions [29].

In total, 20 primer pairs with microsatellites as motifs and nine primer pairs targeting SNPs were tested in the lab.

4.2.1 *In silico* primer tests

The most important criteria for primer selection are listed in section 3.1. To make sure that the criteria were fulfilled, various websites were used. The sequence of the forward primer of primer pair CP8 serves as an example in the following figures.

First, OligoCalc was used to calculate the primer length, as well as the GC content of the primer pairs [58]. Additionally, the melting temperature was noted from the salt-adjusted option (Figure 14).

Oligo Calc: Oligonucleotide Properties Calculator

Enter Oligonucleotide Sequence Below
OD calculations are for single-stranded DNA or RNA

Nucleotide base codes
ACC TGC TTG TTT AGC ACA AT

Reverse Complement Strand(5' to 3') is:
ATT GTG CTA AAC AAG CAG GT

5' modification (if any) 3' modification (if any) Select molecule
50 nM Primer 1 Measured Absorbance at 260 nanometers ssDNA

Calculate Swap Strands BLAST mfold

Physical Constants Melting Temperature (T_m) Calculations

length: 20 Molecular Weight: 6067 GC content: 40%

1 ml of a sol'n with an Absorbance of 1 at 260 nm
is 4.852 microMolar and contains 29.4 micrograms.

Thermodynamic Constants Conditions: 1 M NaCl at 25°C at pH 7.

Rink 33.404 cal/(°K°mol) deltaH 156.9 Kcal/mol
deltaG 24.1 Kcal/mol deltaS 412.1 cal/(°K°mol)

Deprecated Hairpin/self dimerization calculations

5 (Minimum base pairs required for single primer self-dimerization)
4 (Minimum base pairs required for a hairpin)

Check Self-Complementarity

Figure 14: Screenshot of the website Oligo Calc. In the Nucleotide base codes text field, the primer sequence is inserted. The website calculates primer length, GC content and melting temperature. Important values are circled in red.

4.2 In silico primer tests & primer design

To look for secondary structures, the website RNAfold was used [59]. The formation of secondary structures is calculated using the DNA model and tested for 37°C as well as −5°C of the primer-specific melting temperature (Figure 15). Here, the sequence of the primer pairs can be added, and the formation of potential secondary structures is illustrated in the generated graphic (Figure 16). The color ranging from red to green states the probability of occurrence. Green signals that the probability is rather low and red indicates a very high chance of the secondary structure occurring *in vitro*.

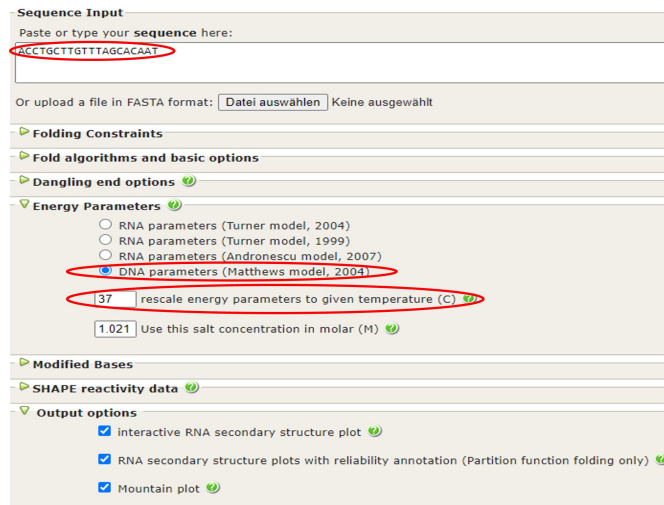


Figure 15: Screenshot of the website RNAfold. In the Sequence Input text field, the primer sequence is inserted. In the section Energy Parameters the DNA model is selected, and the temperature is set to 37°C and −5°C of the primer-specific melting temperature.

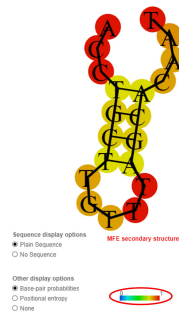


Figure 16: Screenshot of the generated graph by the website RNAfold. The color scale ranging from red to blue, where red equals 1 and blue 0, illustrates the probability of secondary structure formation. Red-colored secondary structures have a high chance of occurring, whereas blue secondary structures have a rather low chance.

4 Experimental part

To make sure, that the primer pair does not form any homo- or hetero-dimers, the website OligoAnalyzer was used [60]. Here, the probability of dimer formation is calculated (Figure 17). The more energy is released during this process the higher the probability of dimer formation (Figures 18 - 20). With an average primer length of 20 bp it is recommended not to use primers which have more than six matching base pairs in the dimer *in silico* test.

OligoAnalyzer

Instructions | Definitions | Feedback

Sequence: 5' MOD INTERNAL 3' MOD MIXED BASES

ACCTGCTTGTTAGACAAT

Bases 20 CLEAR SEQUENCE

Try the new batch mode here

Parameter sets

qPCR (Default)

Target type: DNA

Oligo Conc: 0.2 μ M

Na⁺ Conc: 50 mM

Mg⁺⁺ Conc: 3 mM

dNTPs Conc: 0.8 mM

ANALYZE

HAIRPIN

SELF-DIMER

HETERO-DIMER

NCBI BLAST

TM MISMATCH

ADD TO ORDER

Figure 17: Screenshot of the website OligoAnalyzer. In the Sequence text field, the primer sequence is inserted. With the buttons on the right, the probability of homo and hetero dimer formation can be calculated.

Hetero-Dimer Analysis

Primary Sequence:

5'- ACC TGC TTG TTT AGC ACA AT -3'

Secondary Sequence:

5'- CCGCATAGCAATTTATCTTC -3'

CREATE COMPLEMENT CALCULATE

Figure 18: Screenshot of hetero-dimer-analysis. To analyze the formation of hetero dimers, the sequence of the second primer can be inserted.

Homo-Dimer Analysis

The delta G is calculated by taking into account the longest stretch of complementary bases. These pairs of complementary bases are represented by a solid line. Dotted lines represent additional complementary bases for that dimer structure, but their presence does not impact calculated delta G values. Actual delta G values may vary based on presence of additional complementary bases. The Maximum Delta G value refers to the free energy of the oligo sequence binding to its perfect complement.

Dimer Sequence:

5'- ACCTGCTTGTTTAGCACAAT -3'

Maximum Delta G: -36.2 kcal/mol

Delta G: -6.69 kcal/mol Base Pairs: 4

```
5'  ACCTGCTTGTTTAGCACAAT
      ::::      |||
3'  TAACACGATTGTTTCGTCCA
```

Delta G: -5.24 kcal/mol Base Pairs: 4

```
5'  ACCTGCTTGTTTAGCACAAT
      |||      ::::
3'  TAACACGATTGTTTCGTCCA
```

Delta G: -3.14 kcal/mol Base Pairs: 2

```
5'          ACCTGCTTGTTTAGCACAAT
      ::  ||  ::
3'  TAACACGATTGTTTCGTCCA
```

Figure 19: Screenshot of the result of the homo-dimer analysis. The results are listed ranging from highest to lowest probable dimer formation. Results show the released energy by dimer formation as well as the number of matching base pairs.

Hetero-Dimer Analysis

The delta G is calculated by taking into account the longest stretch of complementary bases. These pairs of complementary bases are represented by a solid line. Dotted lines represent additional complementary bases for that dimer structure, but their presence does not impact calculated delta G values. Actual delta G values may vary based on presence of additional complementary bases. The Maximum Delta G value refers to the free energy of the oligo sequence binding to its perfect complement.

Primary Sequence: 5'- ACCTGCTTGTTTAGCACAAT -3'

Secondary Sequence: 5'- CCGCATAGCAATTTATCTTC -3'

Maximum Delta G: -37.34 kcal/mol

Delta G: -6.69 kcal/mol **Base Pairs:** 4

```
5'      ACCTGCTTGTTTAGCACAAT
      :   |||
3' CTTCTATTTAACGATACGCC
```

Delta G: -5.09 kcal/mol **Base Pairs:** 3

```
5'      ACCTGCTTGTTTAGCACAAT
      :   ||
3' CTTCTATTTAACGATACGCC
```

Delta G: -3.9 kcal/mol **Base Pairs:** 3

```
5'      ACCTGCTTGTTTAGCACAAT
      :   || :  ::
3' CTTCTATTTAACGATACGCC
```

Figure 20: Screenshot of the result of the hetero-dimer analysis. The results are listed ranging from highest to lowest probable dimer formation. Results show the released energy by dimer formation as well as the number of matching base pairs.

Last but not least, the National Center for Biotechnology Information website (NCBI) was used to ensure that the primer binding sites are specific [\[61\]](#). The sequence is added, RefSeq Genome Database is chosen and as species *Cicer arietinum* is added. The Basic Local Alignment Search Tool (BLAST) is set for highly similar sequences (Figure [21](#)). After the search is completed, the table shows the best matching results. The lower the e-value the higher the match. The result shows the specific location of the binding site

4.2 In silico primer tests & primer design

including the chromosome and upper or lower strand (Figures 22, 23). Using the graphics tool, the primer sequence can be searched for as seen in Figures 24 and 25. Following the forward primer sequence, the searched motif, e.g. microsatellites, should be located easily, as well as the sequence of the reverse primer (Figure 26). In order to locate the reverse primer, it has to be transcribed to its reverse complementary sequence 62.

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [?](#) [Clear](#)

ACCTGCTTGTTAGCACAAT

Query subrange [?](#)

From

To

Or, upload file Keine ausgewählt [?](#)

Job Title

Enter a descriptive title for your BLAST search [?](#)

☐ Align two or more sequences [?](#)

Choose Search Set

Database ☒ Standard databases (nr etc.): ☐ rRNA/ITS databases ☐ Genomic + transcript databases ☐ Betacoronavirus

RefSeq Genome Database (refseq_genomes) [?](#)

Organism [Optional](#) ☒ Cicer arietinum (taxid:3827) ☐ exclude [Add organism](#)

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown [?](#)

Exclude [Optional](#) ☐ Models (XM/XP) ☐ Uncultured/environmental sample sequences

Limit to [Optional](#) ☐ Sequences from type material

Entrez Query [Optional](#) [YouTube](#) [Create custom database](#)

Enter an Entrez query to limit search [?](#)

Program Selection

Optimize for ☒ Highly similar sequences (megablast)

☐ More dissimilar sequences (discontiguous megablast)

☐ Somewhat similar sequences (blastn)

Choose a BLAST algorithm [?](#)

BLAST Search database refseq_genomes using Megablast (Optimize for highly similar sequences)

☐ Show results in a new window

Figure 21: Screenshot of Nucleotide BLAST on the NCBI website. Primer sequence is added in the text field in the top left corner. RefSeq Genome Database and organism name (*Cicer arietinum*) are added. The program is selected for highly similar sequences.

4.2 In silico primer tests & primer design

Cicer arietinum cultivar CDC Frontier chromosome Ca4, ASM33114v1, whole genome shotgun sequence

NCBI Reference Sequence: NC_021163.1

GenBank FASTA

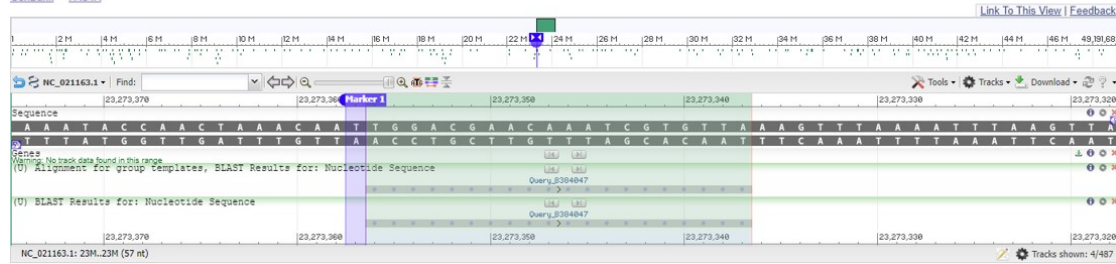


Figure 25: Screenshot of the graphics tool for the highest matching result. The green field shows the primer sequence. A blue marker is placed at the 5' end of the primer sequence.

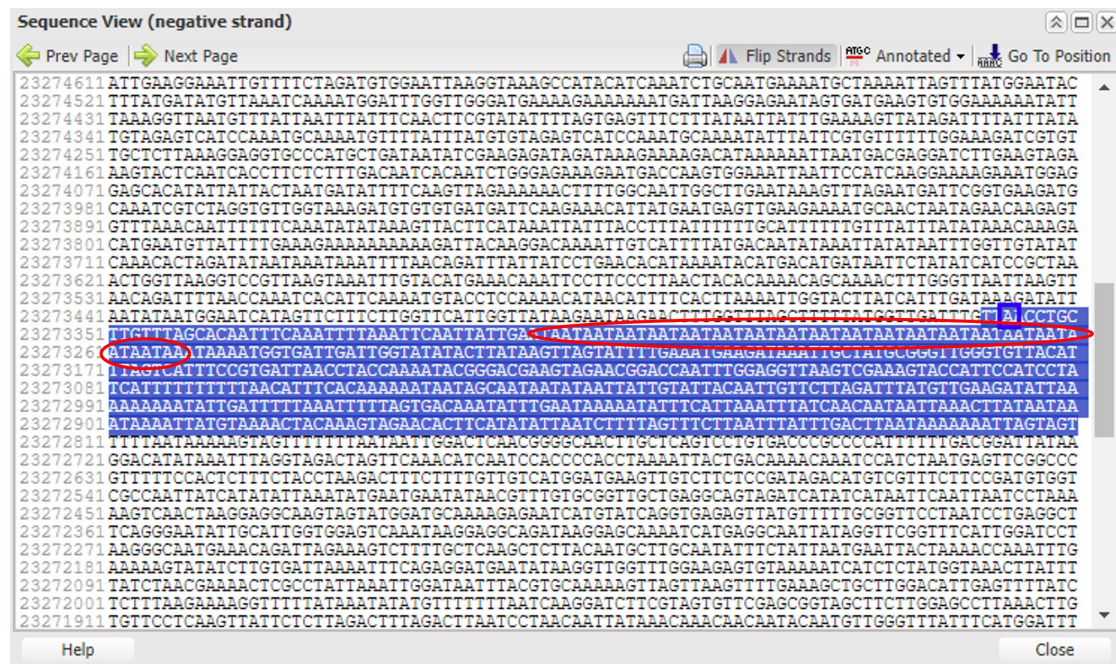


Figure 26: Screenshot of Sequence Text View feature. The microsatellite is circled in red.

An overview of the *in silico* test results of the best primer pair for microsatellites and SNPs can be found in Figures 34 and 58 in section 5.

4.2.2 Optimization

Prior to using the primer pairs for sample differentiation, they were tested in a PCR using the thermocycler QuantStudio™ 5 (Thermo Fisher). For this optimization process, sample CPS1 was used for all primer pairs. This PCR run aimed to find the most suitable annealing temperature and primer concentration for each primer pair. Each primer pair was tested in two concentrations, which most of the time included 0.2 and 0.4 μM and

4 Experimental part

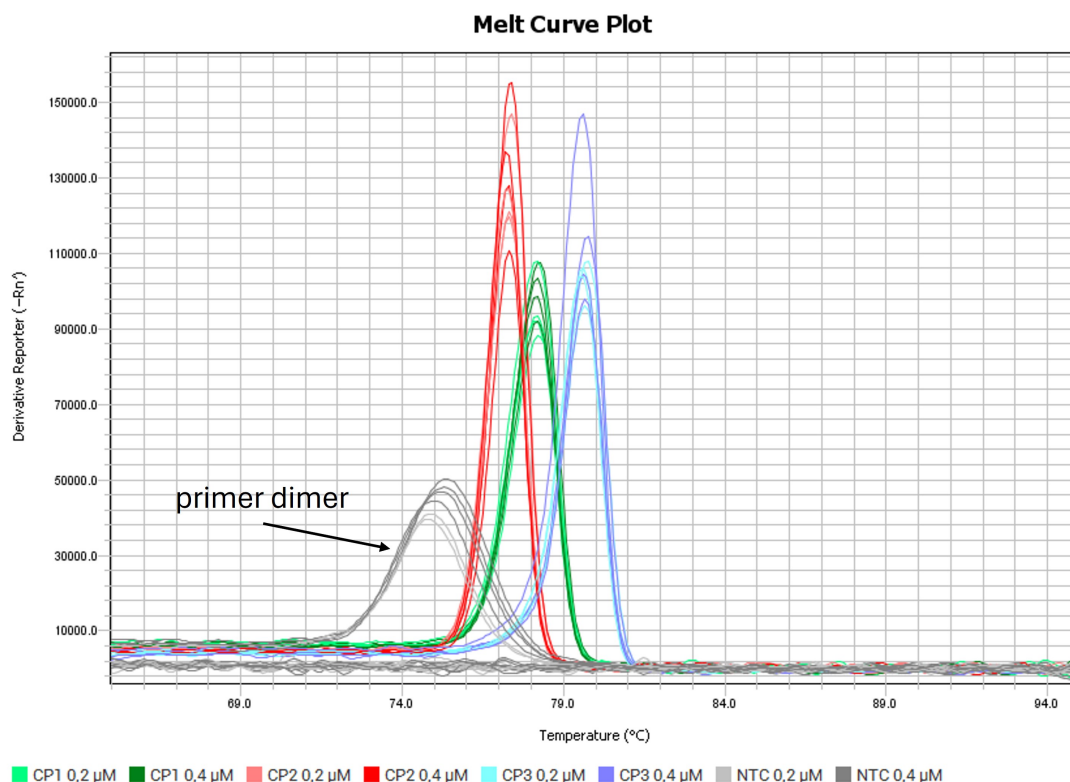


Figure 27: Screenshot of the QuantStudio Design & Analysis Software. Primer pairs CP1 (green), CP2 (red) and CP3 (blue) tested in the concentrations 0.2 and 0.4 μM all generate one singular peak. The no template control (NTC) of primer pair CP2 contains a primer dimer. The dimer peak is marked with an arrow.

rarely ranged up to 0.8 μM . For both concentrations, different annealing temperatures were tested as demonstrated in Table 1. To choose the right temperature range, the melting temperature of each primer pair was taken into account and -1°C and -3°C were deducted. The results were evaluated using the QuantStudio™ Design & Analysis Software by Thermo Fisher. Good results are obtained if there is only one peak per primer pair with high repeatability and a gaussian peak shape. Figure 27 shows an example of the results from three primer pairs CP1, CP2 and CP3 tested in different concentrations (0.2 and 0.4 μM), which all generated one singular peak. The no template control (NTC) of primer pair CP2 contained a primer dimer.

For further information on how the PCR was carried out, check section 4.7

Table 1: Illustration of the metal plate used for QuantStudio™. The two top rows list the used annealing temperature and primer concentration for each column. CP1, CP2 and CP3 are the names of different primer pairs, which are listed in Table 6. NTC serves as the corresponding no template control.

Primer (nM)	53.4		55.4		55.4		57.4		53.3		51.3	
	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4
A	1	CP1	CP1	CP1	CP2	CP2	CP2	CP2	CP3	CP3	CP3	CP3
B	CP1	CP1	CP1	CP1	CP2	CP2	CP2	CP2	CP3	CP3	CP3	CP3
C	CP1 NTC	CP1 NTC	CP1 NTC	CP1 NTC	CP2 NTC	CP2 NTC	CP2 NTC	CP2 NTC	CP3 NTC	CP3 NTC	CP3 NTC	CP3 NTC
D	CP1 NTC	CP1 NTC	CP1 NTC	CP1 NTC	CP2 NTC	CP2 NTC	CP2 NTC	CP2 NTC	CP3 NTC	CP3 NTC	CP3 NTC	CP3 NTC
E												
F												
G												

4.2.3 Primer design

For Primer Design the software PyroMark[®] Q24 Advanced Software was used. New primers were mainly designed for SNPs and not for microsatellites. Additionally, sequencing primers were newly designed for pyrosequencing but the usage of forward primers as sequencing primers provided better results.

First, the DNA sequence containing three or more SNPs was copied into the software and SNPs were chosen as the assay type. When running the program, the software automatically recognized SNPs as the target motif and suggests potential primer sequences. Most of the time, the suggested primer sequences were adjusted manually to further shorten the amplicons. Additionally, the primer pairs were tested *in silico* as described in section [4.2.1](#). If some criteria were not met, the primer sequences were adjusted manually until they were considered suitable. The reverse primer for all SNP primer pairs was biotinylated, as they were scheduled to be sequenced.

Primer pair CP14 was a special case, as it was the only primer pair with a microsatellite motif, which was newly designed. The primer pair sequence taken from the paper published by Sethy *et al.* did not produce repeatable results, but it seemed promising regarding its cultivar differentiation properties. Thus, the primer pair was manually adjusted in two ways:

- Primer pair CP14a: a different location in the genome was chosen for the primer pair sequence resulting in a longer amplicon and the primer sequence was elongated
- Primer pair CP14b: main primer sequence remained the same but was slightly elongated

The improvement of primer pair CP14 is further elaborated in section [5.1.1](#).

4.3 DNA extraction

For DNA extraction the NucleoSpin™ Plant II kit by Macherey-Nagel™ was used (Figure 28). The CTAB and SDS methods are both implemented in this kit. However, the SDS method was used for DNA extraction. For further information on the theoretic background of the SDS method view section 3.2.

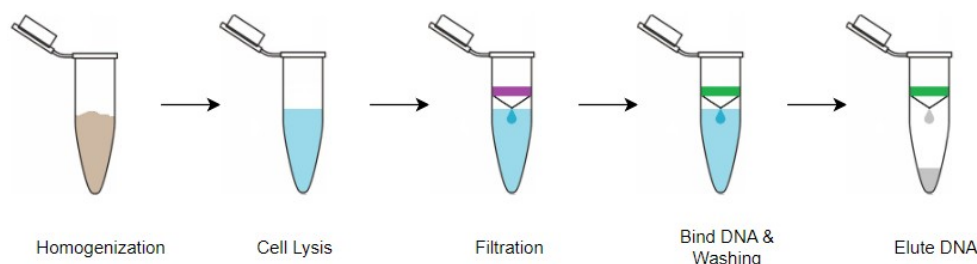


Figure 28: Simplified schematic representation of DNA extraction steps using the NucleoSpin™ Plant II kit by Macherey-Nagel™. Purple and green lines represent two different types of columns used in the kit.

First, the work bench had to be cleaned with DNA exitus in order to prevent contamination. Gloves had to be changed in order to prevent DNA exitus from coming into contact with the samples. The chickpeas were provided as a homogenized powder, and as such, no further sample preparation before the DNA extraction was required.

Approximately 150 mg of the sample were weighed into a 2 mL Eppendorf Tube®. After handling each sample the gloves were changed in order to avoid cross-contamination. It is recommended to work with around eight samples per person per day. DNA was extracted twice from the entire sample set. Thus, the first DNA extraction process for the 24 chickpea samples was carried out on three separate days in October and November. The second DNA extracts of the samples were extracted on two consecutive days together with Klementina Yakova and Sofie Schandl in March. The DNA concentrations of all DNA extracts can be found in the appendix (Tables 10, 11).

900 µL of lysis buffer PL2 and 30 µL of RNase A were added subsequently to each sample. After 15 seconds of pulse vortexing, the samples were shortly spun down with the Galaxy Mini-Centrifuge and briefly vortexed again in order to avoid pellet formation. The samples were incubated on a thermomixer at 65 °C with 600 rpm for 20 minutes. In the meantime, other vials can be prepared. The purple and green NucleoSpin™ filter columns were placed into clean collection tubes, and 5 & 2 mL Eppendorf Tubes® were labeled. For each sample one purple and one green filter column in fresh collection tubes as well as one 5 mL and one 2 mL Eppendorf Tubes® are required.

225 µL of lysis buffer PL3 was added, and after 15 seconds of pulse vortexing, the samples were incubated for five minutes on ice. After the incubation, the samples were centrifuged for one minute at 11,000 rcf. This resulted in the formation of a pellet.

4 Experimental part

The lysate was carefully pipetted into the purple NucleoSpin™ filter column without transferring fragments of the pellet. If the pellet starts to dissolve or is not very stable, it is recommended to repeat the prior centrifugation step to clear the lysate. After a successful transfer, the lids of the filter columns were carefully closed in order to avoid cross-contamination from potential splashing droplets, and the samples were vortexed for two minutes at 11,000 rcf. If there are any remnants of the lysate, which have not passed the filter, this step is to be repeated.

The clear filtrate was collected in a 5 mL Eppendorf Tube®. 1350 µL of the PC buffer were added to each sample, followed by 15 seconds of pulse vortexing. Then, 675 µL of the sample were carefully pipetted into the green NucleoSpin™ filter column. It is important to apply no more than 700 µL to the column, as it would exceed its maximum capacity. The lids of the filter columns were carefully closed, and the samples were centrifuged for one minute at 11,000 rcf. The filtrate was then discarded. This step was repeated until the entire sample contained in the 5 mL Eppendorf Tube® was applied to the column.

Next, the columns were washed with 400 µL PW1 buffer. The lids of the columns were closed, and the samples were centrifuged for one minute at 11,000 rcf. 700 µL of PW2 buffer were added to each column, and the centrifugation step is repeated with the same settings. As a last washing step, 200 µL of PW2 buffer were added, and the steps above were repeated. However, this time the centrifugation was elongated to two minutes. After each washing step, the filtrate was discarded.

The green filter column was placed in a new 2 mL Eppendorf Tube®, which had been labeled prior. 50 µL of the PE buffer, which was preheated to 65 °C, was added to each column. It is important to pipette the PE buffer in the centre of the filter column in order to ensure a better distribution. The lids of the columns were closed, and the samples were incubated for five minutes on a thermomixer set to 65 °C. After the incubation, the samples were centrifuged for one minute at 11,000 rcf. Thus, the first elution (E1) was obtained.

If necessary, the filter column can be placed in another new 2 mL Eppendorf Tube®, and the application of the 50 µL PE buffer can be repeated in order to obtain a second elution (E2).

The DNA extracts E1 and E2 have to be stored at −20 °C.

4.4 Fluorimetry

In order to measure the concentration of the obtained DNA the Qubit™ 1x dsDNA BR Assay Kit was used. The amount of solutions required depends on the number of samples, which are to be measured. Thus, it is recommended to use the solution calculator provided by the Qubit™ 4 Fluorometer in order to obtain the required volumes of buffer solution and color reagent. After combining both of the solutions, the mixture was vortexed for five seconds.

For each sample and for the two standards Qubit™ Assay Tubes were prepared. In order to make labelling and pipetting easier, the tubes were placed into 2 mL Eppendorf Tubes®, whose lids had been removed. This also served as a precautionary step, as the

bottom half of the Qubit™ Assay Tubes are not supposed to be touched.

For the standards 190 μL and for the samples 198 μL of the buffer-colorant solutions is required. First, 10 μL of the two standard solutions, which are included in the kit, were added to the corresponding Qubit™ Assay Tubes. After closing the lids, the tubes were vortexed for three seconds and incubated for two minutes at room temperature. After the incubation, the standards were ready for measurement. It is important to choose the right settings for the measurement with the Qubit™ 4 Fluorometer, which include double-stranded DNA and broad range kit. After measuring the two standards, a calibration curve is generated. In comparison to the standards, only 2 μL of the samples were needed for the concentration measurement. The procedure for the standards was also applied to the samples. For a large number of samples, it is recommended to measure the DNA in several runs, as the mixture of DNA and buffer-colorant solution is not very stable for a prolonged time.

4.5 Gel electrophoresis

In order to prepare one gel, around 3 g of agarose powder were weighed into an erlenmeyer flask. 10xTBE buffer was diluted to 0.5xTBE by taking 50 mL of the 10xTBE buffer and filling it up to 1 L with distilled water. 100 mL of the 0.5xTBE buffer were then used to dissolve the agarose powder. The erlenmeyer flask was placed on a heating plate, covered with a watch glass and heated to approximately 180 °C for around 30 minutes. The agarose solution is ready once it has been completely dissolved.

In the meantime, the gelcaster was prepared. A spirit level was used to make sure that the gelcaster is evened out. Depending on how many samples are to be applied on the gel either the comb forming 26 or 30 pockets can be used.

Once the agarose solution was ready, it had to be poured quickly into the gelcaster. Waiting too long for the transfer might cause the agarose to cool down and solidify in the erlenmeyer flask. Potential air bubbles which might appear after the transfer can be removed with pipette tips or moved to the bottom corners of the gelcaster. The gel takes around 45 minutes to solidify.

In the meantime, the loading buffer can be prepared and the samples, which are to be analyzed, can be thawed. For a gel containing 28 samples and two markers the following volumes of reagents listed in Table 2 were necessary:

Table 2: Used recipe for gel loading buffer.

Loading buffer	V (μL)
Nuclease-free water	71.9
6X LD orange	35.9
Total volume	107.8

3.5 μL of the well-mixed loading buffer were pipetted in pre-labeled 200 μL Eppendorf Tubes®. Next, 3.5 μL of the corresponding DNA amplicons were added. By pipetting

4 Experimental part

up and down repeatedly, the solution is mixed. Afterwards, the amplicons can be frozen again. The marker solution was mixed with the solutions listed in Table 3 below:

Table 3: Used recipe for DNA ladder (marker).

Marker (DNA ladder)	V (μL)
Nuclease-free water	10.5
Ladder	0.6
6X LD orange	2.2
Total volume	13.3

Once the gel is hardened, the comb can be slowly removed in order to not damage the pockets. The gel was removed from the gelcaster and placed into the gel running chamber, which was then filled with 0.5xTBE until the gel was completely covered. When placing the gel, it is important to keep the running direction of the gel in mind. 6 μL of the sample solutions can be loaded into the pockets. The first and the last pocket of the gel are reserved for the DNA marker. After the loading of the gel was completed, the lid of the gel chamber was placed on top and the chamber was connected to the power supply. The gel ran at 80 mV for around two hours. It is recommended to take a look at the gel once in a while to make sure how far the sample front has proceeded. Once the front was around two centimetres away from the bottom end of the gel, the run was stopped.

The gel was removed from the chamber and placed into a tupperware box. The 0.5xTBE solution remaining in the chamber was poured into the drain, and the chamber was rinsed with distilled water, dried and stored. The gel was covered in dying solution consisting of 150 mL distilled water and 45 μL 1000xGelRed[®] Nucleic Acid Stain. This solution can be used to stain three gels. Afterwards, it has to be discarded, and a new dying solution has to be prepared. As the solution is very sensitive to light, the bottle has to be covered with aluminium foil and stored in a dark place. Once the gel was covered, the lid of the tupperware was placed on top, and the gel was incubated in a dark place for 30 minutes.

After staining the gel, the dying solution can be poured back into the storage bottle. The gel was covered with distilled water to remove any excess dye and incubated for another 30 minutes in a dark place. The gel was then removed from the tupperware box and placed on the Herolab Transilluminator. When placing the gel, it is important to avoid trapping air bubbles underneath it. The light in the room was switched off, and the UV light (312 nm) of the Transilluminator can be turned on, making the DNA bands visible.

4.6 Primer dilution & aliquotation

Primers were delivered as lyophilized pellets. First, the pellets were relocated to the bottom of the vials by shaking them or by brief centrifugation using the Galaxy Mini-Centrifuge. The primer dilution and aliquotation were carried out in the Gerner lab on a sterile bench. The specific volumes of RNase free water, which are needed to dilute the

primers to a concentration of 100 μM are stated on the technical datasheet, which was delivered together with the primer pairs. After adding the correct amount of RNase free water, the vials containing the forward and reverse primers were vortexed three times for ten seconds every five minutes for a total of 20 minutes. The primer pairs were aliquoted in accordingly labeled 200 μL Eppendorf Tubes[®], each containing 10 μL of primer solutions. For forward primers, green-colored and for reverse primers, yellow-colored Eppendorf Tubes[®] were used.

Only one primer pair at a time was diluted, and after each primer pair, the gloves were changed in order to avoid cross contamination. The diluted primers were stored at -20°C .

4.7 Polymerase chain reaction

The preparations for PCR were carried out in three separate labs to avoid potential contamination. Additionally, for each lab a different lab coat was used. The samples, primer pairs, master mix and a metal plate to transport and cool the aliquoted master mix are stored in the Cichna lab at -20°C . The primer pair, nuclease-free water, 2x Type-it[®] HRM PCR master mix, a cool bag and the metal plate were put into a styrofoam box to make transportation to the other labs easier.

In the Gerner lab, the master mix was prepared. Prior, the required volumes of all components are calculated. For a standard PCR using one primer pair and all chickpea samples, which were analyzed in duplicates, including a no template control (NTC), 52 approaches are required (view Table 4).

Table 4: Standard master mix recipe used for chickpeas with a primer concentration of 0.2 and 0.4 μM . The scale is set to 52 samples (25 chickpea samples and one NTC in duplicates). Fw and rev stand for forward and reverse primer respectively.

PCR master mix		
Solution	0.2 μM	0.4 μM
Nuclease-free water	412 μL	366 μL
2x Type-it [®] HRM PCR master mix	572 μL	572 μL
fw primer (10 μM)	23 μL	45.8 μL
rev primer (10 μM)	23 μL	45.8 μL
Total volume	1030 μL	1029.6 μL
Master mix per well	18 μL	18 μL
DNA extract	2 μL	2 μL

After all reagents were thawed, the primer pair was diluted to 10 μM by adding 90 μL nuclease-free water to the aliquots. The primer pair was then vortexed three times for at least ten seconds at the highest setting. Additionally, 10 μL of nuclease-free water were aliquoted in a new 200 μL Eppendorf Tube[®], which served as the NTC.

4 Experimental part

A new 2 mL Eppendorf Tube[®] was required for the master mix. First, the correct amount of nuclease-free water was pipetted into the Eppendorf Tube[®]. Before pipetting the master mix, it was pulse vortexed three times for around five seconds at lower settings, as too rough vortexing might damage the polymerase. The corresponding volume of the master mix was pipetted into the 2 mL Eppendorf Tube[®] followed by the forward and the reverse primer. The master mix solution is vortexed three times for five seconds with the lower setting. 18 μ L of the master mix solution were pipetted into each Strip Tube and closed with Strip Caps, which were labeled with the numbers from 1 - 52. The reaction vials were placed into the metal box for easier transportation and to keep them cooled.

Primers, master mix and nuclease-free water were returned to the Cichna lab, and the DNA extracts of the samples were thawed and properly vortexed. They were placed in the styrofoam box and carried to the Marko lab.

In preparation for the PCR, the DNA extracts were diluted to 1:100 with nuclease-free water by first pipetting 198 μ L of nuclease-free water into 2 mL Eppendorf Tubes[®] in the Gerner lab. In the Marko lab 2 μ L of the corresponding DNA extract were added to the Eppendorf Tube[®]. The total volume of the DNA dilution is 200 μ L, which suffices for 50 PCRs.

In the Marko lab, 2 μ L of the diluted DNA samples were added to the corresponding Strip Tubes. The last two wells are reserved for the NTC and therefore 2 μ L of nuclease-free water was added. Afterwards, the Strip Tubes were closed again with the Strip Caps.

The Rotor-Gene Q thermocycler is located inside a separate room in the Gerner lab. The samples were placed inside the sample rotor so that the numbers of the rotor aligned with the sample numbers. The PCR run can be set up using Rotor-Gene Q Series Software. As explained in section 3.5 PCR consists of multiple steps requiring different temperatures. The temperature program has to be adjusted accordingly (view Table 5). Under the samples tab, the sample names can be added, and for each sample, a different color can be selected. After checking that all parameters were set correctly, the PCR run can be started. The samples, the cool bag and the metal block were put back in the freezer.

Table 5: Overview of the temperature program for PCR and HRM. Annealing temperature depends on the melting temperature of the used primer pair.

	Step	Time	Temperature ($^{\circ}$ C)	Repeat
Hold 1	denaturation, activation of polymerase	5 min	95	
Cycling	denaturation	15 s	94	
	annealing	30 s	variable	50 cycles
	elongation	30 s	72	
Hold 2	final elongation	10 min	72	
Hold 3	denaturation	1 min	95	
Hold 4	strand pairing	1 min	40	
HRM	HRM	-	62 - 95	

After the PCR was complete the Rotor-Gene Q thermocycler performed the HRM.

4.7.1 Optimization

The elongation temperature was changed for microsatellites composed of only A and T. This includes microsatellites targeted by the primer pairs CP5, CP6, CP7, CP8, CP10, CP11, CP15, CP17, CP18, CP19 and CP20. A and T only have two hydrogen bonds and an elongation temperature of 72 °C was too high for them to be stable and produced results as listed in Figure 45. So 68 °C was used instead for the elongation and final elongation steps, which led to repeatable results.

4.8 High resolution melting analysis

HRM was carried out by the Rotor-Gene Q thermocycler right after the completion of 50 PCR cycles. During HRM the temperature was gradually increased from 62 °C to 95 °C with a rate of 0.05 °C/s. After the run was completed, the amplicons were stored in a freezer at −20 °C for further experiments.

To evaluate the results, the Rotor-Gene Q Series Software was used. When opening the finished run, the amplification curves are shown (Figure 29).

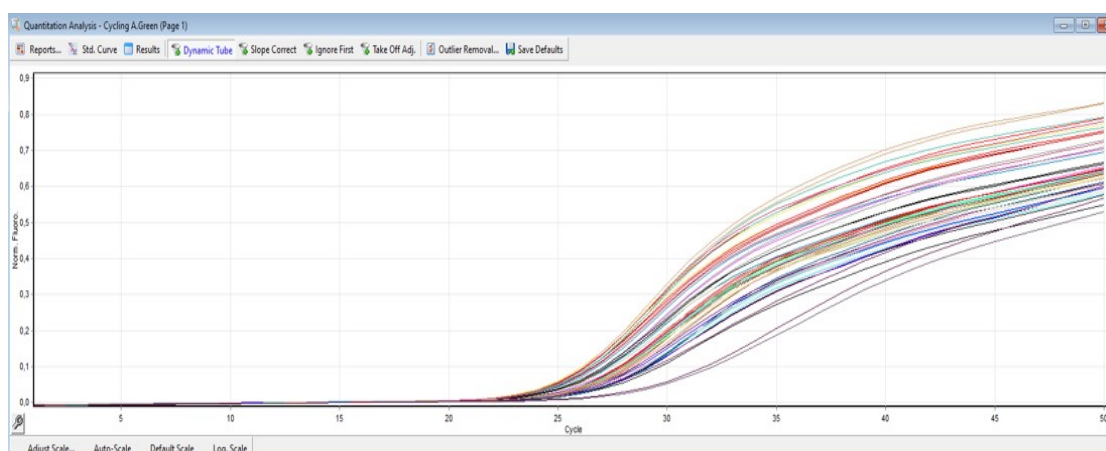


Figure 29: Screenshot of the Rotor-Gene Q Series Software of the amplification curve of all samples with primer pair CP8.

Under the tab analysis HRM was selected. First, the curves were normalized. Therefore the boundaries were set to regions, where all the curves are still highly identical. The border encloses the region in which the melting curves are highly variable (Figure 30). The distance between the start and end points of the leading and trailing range was set to 1.5. Reasons, why normalization is important, can be found in section 3.6.

4 Experimental part

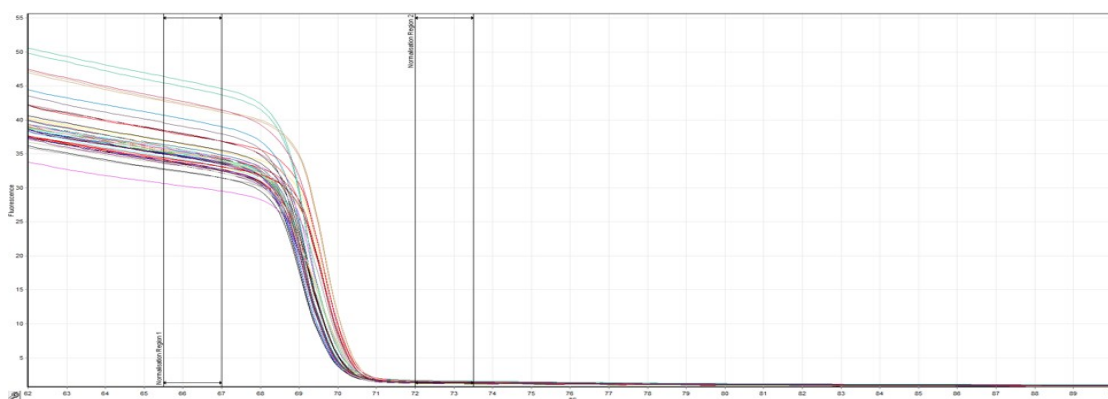


Figure 30: Screenshot of the Rotor-Gene Q Series Software during the normalization step of primer pair CP8. The borders can be adjusted manually.

After normalization, the repeatability of the results was checked by comparing curves that were obtained for one and the same sample analyzed in duplicate. For most of the tested primer pairs a high repeatability could be achieved. Exceptions are primer pair CP14 as well as primer pair SNP7, which is further discussed in section 5.

To evaluate if the primer pair was suitable for cultivar differentiation, the melting curves of each sample were compared with one another, and it was decided whether the samples could not be differentiated (red), could probably be differentiated (yellow) or could reliably be differentiated (green). As an overview, an excel data-sheet is created for each primer pair, where the color code above is used to assign each sample pair a property (differentiated - not differentiated) (Figure 31).

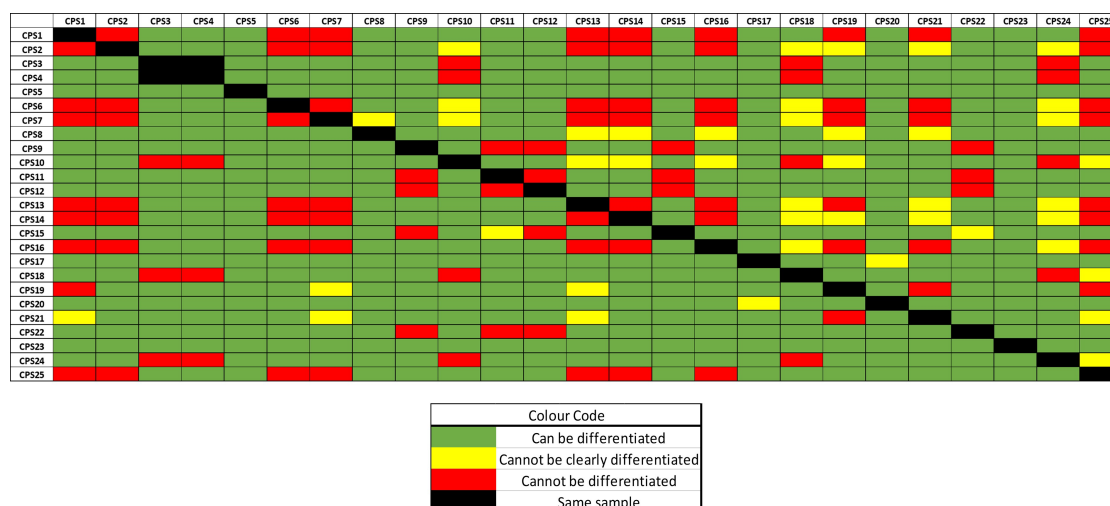


Figure 31: First evaluation draft for primer pair CP8. The first row and column list samples 1 to 25 in ascending order.

To decide which samples are differentiable and which are not, a temperature difference of at least 0.1 °C was set as the lower limit to mark samples as differentiated. Prior primer pairs with promising results were reevaluated (Figure 32). This process is further explained in the section 5.1.2. Samples whose melting curves are very close to each other but could theoretically be differentiated from our obtained results were marked as red (not differentiable) to account for possible scattering, which might happen if the experiment is repeated in a different lab under different circumstances. Thus, only samples whose melting curves were highly distinguishable were marked as green.

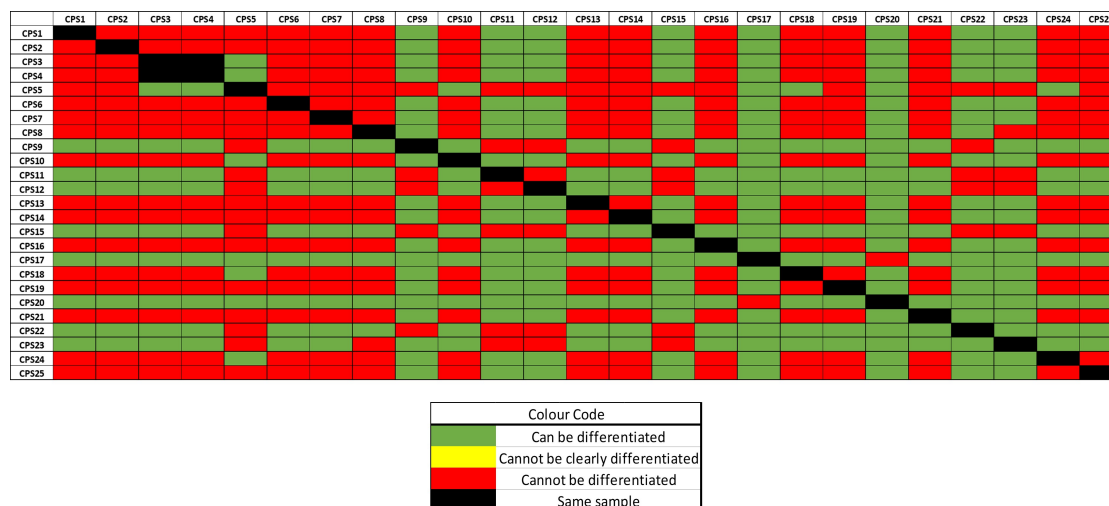


Figure 32: Second evaluation of primer pair CP8. The first row and column list samples 1 to 25 in ascending order.

For each primer pair one excel data-sheet was created. Thus, 22 excel sheets for microsatellites and nine for SNPs, which sums up to a total of 31 excel sheets were created. As a final step, individual primer pairs were compared to each other and complementary primer pairs were grouped together to create a final excel sheet for microsatellites and SNPs by using as few primer pairs as possible to differentiate most cultivars (Figures 50, 61).

4.9 Pyrosequencing

The sequencing run was prepared using the PyroMark[®] Q48 Autoprep software. Here, the sequence, which is to be sequenced, was added beforehand, and the location of the SNPs was indicated with a backslash.

First, the sequencing primers were prepared. Therefore, one of the 100 µM aliquoted forward primers was diluted with 90 µL ELGA water to 10 µM. From this dilution a 4 µM dilution was prepared by taking 97.6 µL of the 10 µM dilution and adding it to 146.4 µL ELGA water. The primer dilution was properly vortexed. The PSQ kit, the diluted primer, a cool bag, and the thawed amplicons were taken to a different room of the Marko

4 Experimental part

lab for sequencing. A cool bag was required to keep the enzyme and substrate cooled.

Sequencing was performed with the PyroMark[®]Q48. Before usage, an absorber strip was inserted and all channels were rinsed with ELGA water twice during the implemented cleaning process. In order to start sequencing the directions by the PyroMark[®]Q48 device had to be followed. The channels were loaded with corresponding volumes of the dNTPs, buffer solutions, as well as enzyme and substrate, which were all included in the PSQ kit. The primer was also loaded in the signified channel. 3 μ L of magnetic beads and 10 μ L of the corresponding DNA template were loaded into each well of the sample disc and then the sequencing was started.

After the run was completed, all channels were rinsed again with ELGA water, and the sample disc and the absorber strip were removed from the device.

To evaluate the sequencing result, the run file was opened with the PyroMark[®] Q48 Autoprep software. The accuracy of the sequencing result was color-coded. Blue boxes stand for correctly determined SNPs, yellow is still reliable but indicates some form of mismatch and red signals, that the result is not very reliable as there might have occurred some form of error (Figure 33).

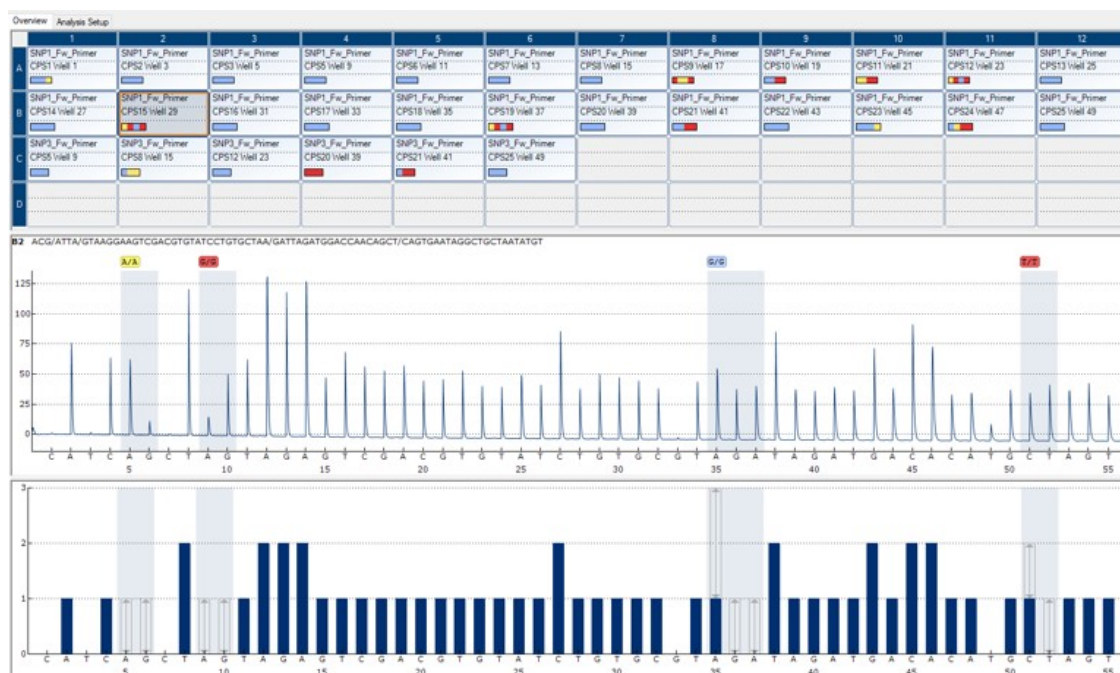


Figure 33: Screenshot of the result of a sequencing run depicted in the PyroMark Q48 Autoprep software. The picked result shows a sequence containing all three listed color codes. The upper section of the screenshot shows the analyzed samples including the primer pair and the well number. The middle section lists the recorded sequence of the selected samples. Each peak represents an inserted nucleotide, and the intensity of said peak is proportional to the number of inserted nucleotides of the same type. The gray fields indicate the locations of the SNPs. The bottom section is a histogram representing the predetermined sequence, which is selected prior to the sequencing run. The gray areas represent the SNP locations and indicate possible sequencing outcomes.

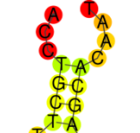
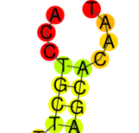
5 Results & discussion

The discussion is entirely based on the interpretation of the received results as so far there are no similar publications to be found.

5.1 Microsatellites

5.1.1 Suitability of primer pairs

Multiple primer pairs targeting microsatellites were tested *in silico* and filtered for specific criteria as listed in sections 3.1 and 4.2. A summary of the *in silico* test results of primer pair CP8, which was the most promising regarding its cultivar differentiation ability, is listed in Figure 34.

CP8	Primer sequence	Length (bp)	GC content (%)	TM (°C)	Rep
Fw	ACCTGCTGTTGTTAGACAAT	20	40	54.3	(TAA)
Rev	CCGCATAGCAATTTATCTTC	20	40	54.3	
Sec	 TM -5 °C (Forward)	 TM 37 °C (Forward)	 TM -5 °C (Reverse)	 TM 37 °C (Reverse)	
Dim	Delta G: -6.69 kcal/mol Base Pairs: 4 5' ACCTGCTGTTGTTAGACAAT 3' TAACGATTTGTTGTCGCA Delta G: -5.24 kcal/mol Base Pairs: 4 5' ACCTGCTGTTGTTAGACAAT 2:2:2 3' TAACGATTTGTTGTCGCA Forward-Dimers	Delta G: -5.36 kcal/mol Base Pairs: 4 5' CCGCATAGCAATTTATCTTC : : : : : : 3' CTCTATTTAACGATACGCC Reverse-Dimers	Delta G: -6.69 kcal/mol Base Pairs: 4 5' CCGCATAGCAATTTATCTTC : 3' TAACGATTTGTTGTCGCA Delta G: -5.09 kcal/mol Base Pairs: 3 5' CCGCATAGCAATTTATCTTC 3' TAACGATTTGTTGTCGCA Forward + Reverse-Dimers	Amplicon length: 172 bp	

ACCTGCTGTTTAGCAAAATTCAAATTTTAAATTC AATTATTGAAATAATAATAATAATAATAATAATA
TAATAATAATAATAATAATAATAATAAATAAAAAATGGTGATTGATTGGTATACTTATAAGTTAGTATTTTGAAA

Figure 34: Overview of the *in silico* test results of primer pair CP8 including primer sequence, length, GC content, melting temperature (T_M), repetitive motif (Rep), amplicon length, secondary structures (Sec) and potentially formed dimers (Dim). At the bottom, the location of the forward and reverse primer and the microsatellite are marked with yellow and cyan, respectively.

20 different primer pairs were tested in the lab according to sections 4.7 and 4.8. The sequences of the primer pairs, the used experimental conditions, and their source are listed in Table 6. New names were assigned to the primer pairs including the abbreviation CP (chickpea primer) and numbers 1 to 20. Additionally, the table lists the potential of the singular primer pairs to differentiate the tested chickpea cultivars based on the results obtained in this thesis.

Table 6: All primer pairs tested for cultivar differentiation targeting microsatellites in this thesis including their original name and source as well as their sequence, their used concentration (Conc.), annealing temperature (T_A) and whether they were suitable for differentiation (Y) or not (N). For primer pairs marked as (N) the melting curves of all cultivars were overlapping.

Original name	New name	fw-sequence	rev-sequence	Conc. (μ M)	T_A ($^{\circ}$ C)	Source	Differentiation
NCPGR21	CP1	TCTACCTCGTTTTTCGTGCC	TTGCTCCTTCAGAAAACCC	0.2	55.4	Sethy <i>et al.</i>	Y
NCPGR33	CP2	ACATCTTGAAGTGGCCCAAC	TGCAAGCAGAGGGTTAGACG	0.2	55.4	Sethy <i>et al.</i>	Y
NCPGR99	CP3	ATCATGAAGCAATCGTCAC	TGAACCCAAACATAGCATACA	0.2	53.3	Sethy <i>et al.</i>	N
NCPGR72	CP4	TTAACCACATTAGCGTGACTT	GATGAGCTTCTGCTTTCAT	0.2	53.3	Sethy <i>et al.</i>	Y
TA78	CP5	CGTAATAATAAGTTTCCCTCC	CATCGTGAATATTGAAGGGT	0.2	53.3	Winter <i>et al.</i>	Y
TA176	CP6	ATTTGGCTTAACCCCTCTTC	TTTATGCTTCTCTCTCTCG	0.4	53.3	Winter <i>et al.</i>	Y
TA180	CP7	CATCGTGAATATTGAAGGGT	CGGTAATAAGTTTCCCTCC	0.2	53.3	Winter <i>et al.</i>	Y
TR20	CP8	ACCTGCTTGTTTAGCAGCAAT	CCGCATAGCAATTATCTTC	0.4	53.3	Winter <i>et al.</i>	Y
GA2	CP9	TGCATTTGGAATAATCAGCATGA	AATTTGGTTTCGCACAAC	0.2	53.3	Winter <i>et al.</i>	N
GAA44	CP10	AGCAAGCCCATGATTTCTC	ATGACATTCGAAATGGGCTTC	0.4	55.4	Winter <i>et al.</i>	N
GAA54	CP11	TGACGAGAGATTAAGAGAGAA	TGGAGGCAGTTCAAGTTTG	0.2	35.4	Winter <i>et al.</i>	Y
NCPGR51	CP12	CATATGCAAGGGCAATTAG	CTCTATCTTCATGTTGGCG	0.2	53.3	Sethy <i>et al.</i>	N
NCPGR55	CP13	TCCATTGGATACATCACAGG	GGCGAAATTCAGTATTTTGG	0.2	51.3	Sethy <i>et al.</i>	N
NCPGR90	CP14	TAGCATACGATTTGTCAAGCA	AAGAGCAGATAGGGTTTGT	0.4	51.3	Sethy <i>et al.</i>	Y
	CP14a	AACTTGGCAACGAATCAGTCAG	GGCAAGGAAGGATCAATAACA	0.4	57.4		Y
	CP14b	GAGGTAGCATACCATTTCAACCA	CTAAGAGCACATACGGTTTGTGT	0.4	61.0		Y
TA11	CP15	CATGCCATAAAGCTCATACATACAAAC	TTCAATTGAGGACAAATGTAATTTAG	0.2	57.8	Winter <i>et al.</i>	N
TA13	CP16	TAAGTTAAGGGAGCAAGCAA	CAAGTTGAGTGTAAACCAAT	/	/	Winter <i>et al.</i>	N
TA21	CP17	GTACCTCGAAGATGTAGCCGATA	TTTTCATTTAGAGTAGCATCTCTTG	0.4	61.6	Winter <i>et al.</i>	Y
TA43	CP18	GGTTGTGTTCTCCAGATTTT	AAGAGTTGTTGGAGAGCAA	0.4	50.0	Winter <i>et al.</i>	N
TA106	CP19	CGGATGGAAGTCAACTTATC	TGCTGATGATGTGATCTGT	0.4	53.3	Winter <i>et al.</i>	Y
TS72	CP20	CAACAATFACATAAAGTATTTCTCT	AAAAATTGATGGAACAAGTGTATTATG	0.4	58.2	Winter <i>et al.</i>	Y

Primer pair CP14 seemed promising for cultivar differentiation but the amplification curves were not repeatable (Figure 35). The product obtained with CP14 produced two melt peaks with varying intensity between cultivars, which suggests that two different products of varying sizes were amplified. It was assumed, that the high cycle number might have led to the formation of artefacts, and thus in a second attempt a few samples whose amplicons produced melting curves which either showed very good or no repeatability were handpicked and tested again under the same conditions but with a reduction of the cycle number from 50 to 43 (Figure 36). While the scattering of the melting curves of all tested samples with low repeatability was reduced, the products of samples 14 and 21, which showed good repeatability in the prior experiment, were now scattering. This led to the assumption that the primer pair might bind to a second unspecific position with varying probability and two products are amplified. Products that showed good repeatability, such as for samples 13, 14, and 23, only had one melt peak, whereas products of samples with bad repeatability had two. The second peak might belong to the unwanted product.

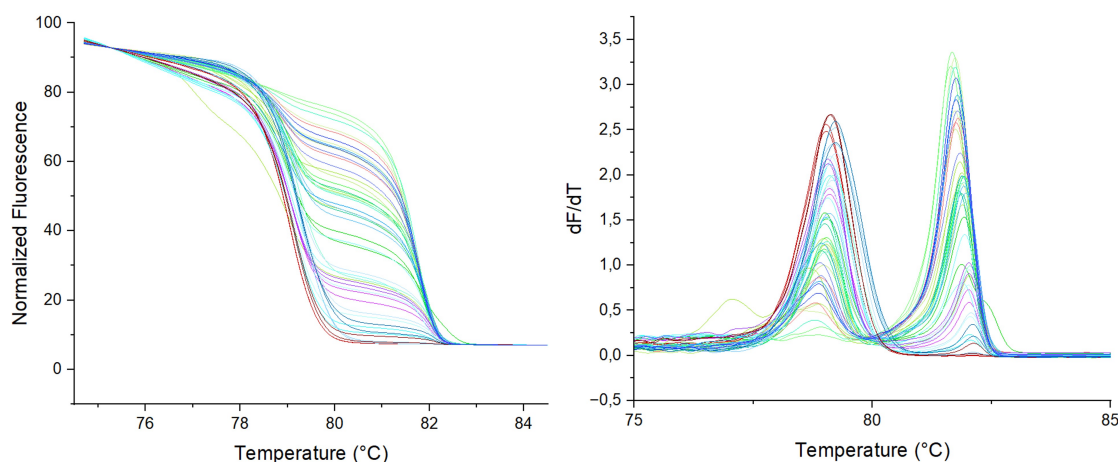


Figure 35: The left side shows the melting curves of the amplicons of all 25 chickpea samples analyzed in duplicates using primer pair CP14. Identical samples share the same color. The section between 79 and 83 °C seemed promising for cultivar differentiation, but the repeatability of the results was rather low. On the right the two melt peaks are depicted. Varying peak intensities between the different samples can be observed.

5 Results & discussion

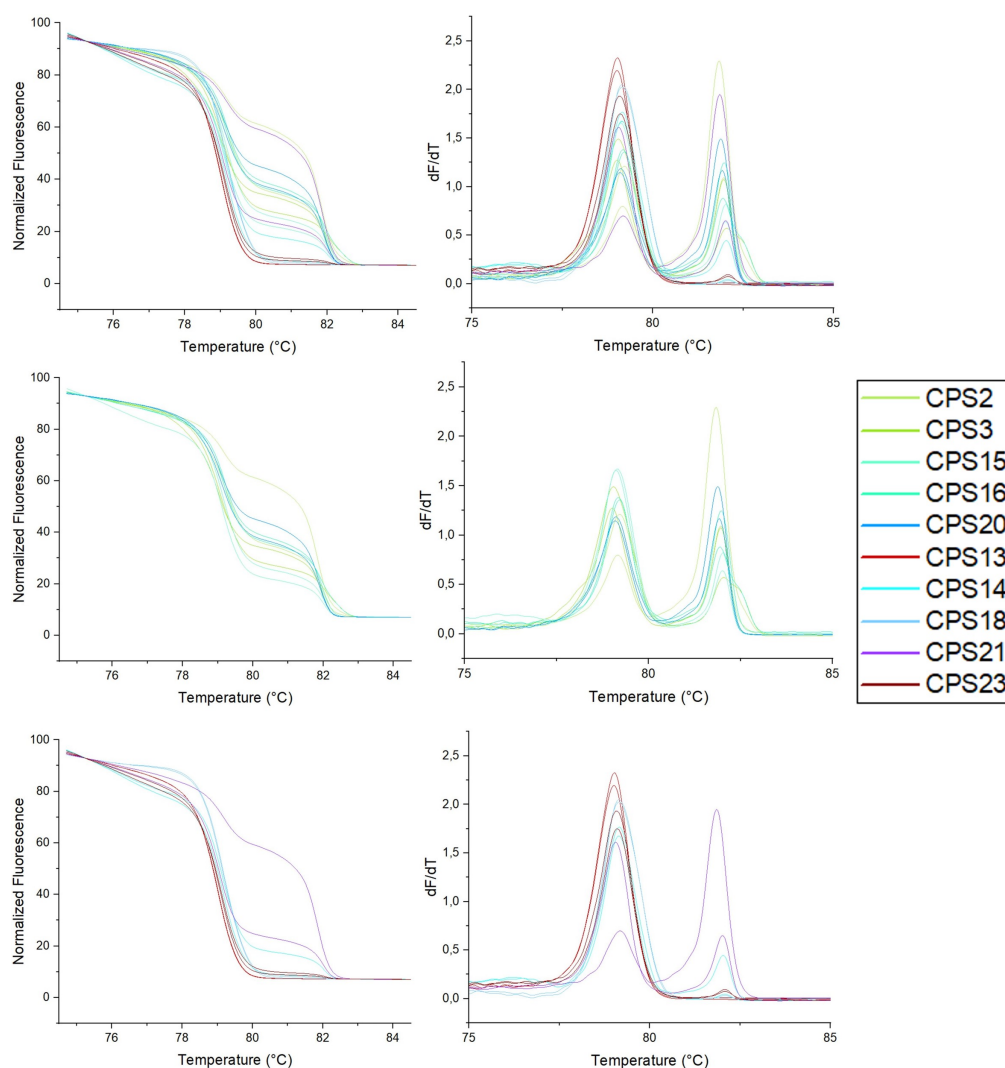


Figure 36: At the top, all the melting curves (left) and melt peaks (right) of the products of the handpicked samples analyzed in duplicates using primer pair CP14 are depicted. Identical samples share the same color. The middle row shows the melting curves (left) and melt peaks (right) of the products of samples 2, 3, 15, 16, and 20, which showed low repeatability in the prior experiment. While the scattering was slightly reduced, it can still be observed in this graphic. The bottom row shows melting curves (left) and melt peaks (right) of the products of samples 13, 14, 18, 21 and 23, which showed good repeatability in the prior experiment. While 13 and 18 remained the same, 24, 21 and 23 were scattered.

To increase the primer binding specificity the primer sequence was redesigned twice as explained in section [4.2.3](#):

- Primer pair CP14a: a different location in the genome was chosen for the primer pair sequence resulting in a longer amplicon, and the primer sequence was elongated
- Primer pair CP14b: main primer sequence remained the same but was elongated

Both new primer sequences were tested for their differentiation ability with all chickpea samples. While the repeatability slightly increased for primer pair CP14b, it did not significantly improve for primer pair CP14a. However, different melting behaviour was observed. While for 14a both melt peaks were still present with varying intensity (Figure 37), 14b only showed one remaining melt peak (Figure 38). Additionally, the amplification with primer pair CP14a did not work properly for some samples as they showed a high CT-value (Figure 39). Especially CPS6 was not well amplified. This might be explained by the high tendency of the forward primer of CP14a to form a hairpin secondary structure.

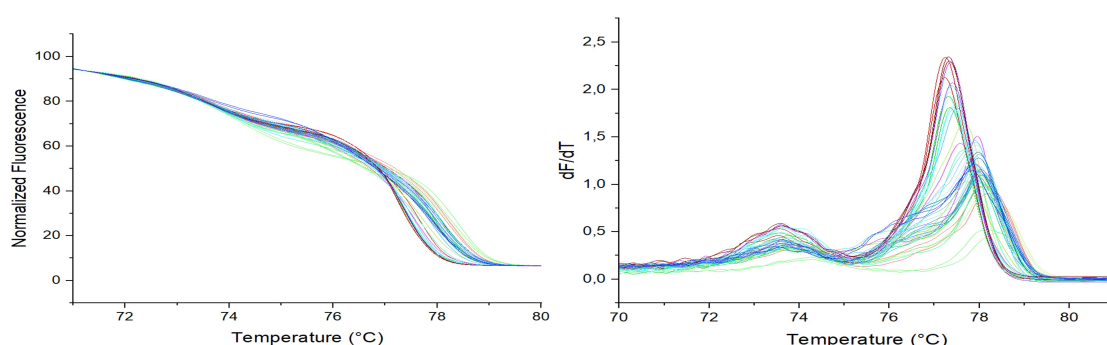


Figure 37: The left side shows the melting curves of the products of all 25 chickpea samples analyzed in duplicates using primer pair CP14a. Identical samples share the same color. While the repeatability could still be improved, the melting curve is more complex compared to CP14b and might provide more potential for cultivar differentiation. On the right the two melt peaks are depicted. Varying peak intensity between the products of the different samples and a shift for the peak maximum of the second peak between individual samples can be observed.

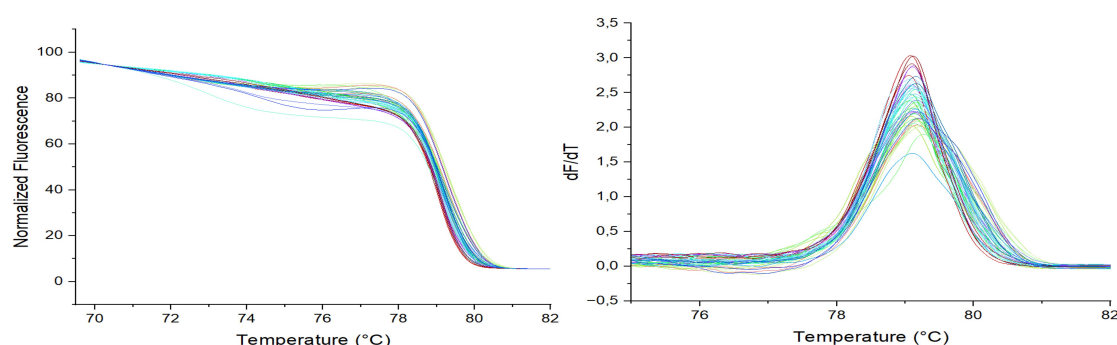


Figure 38: The left side shows the melting curves of the products of all 25 chickpea samples analyzed in duplicates using primer pair CP14b. Identical samples share the same color. While the repeatability increased slightly in comparison to primer pair CP14, the melting curves of the products of all samples look highly similar. On the right the melt peak is depicted. The peak intensity between the products of all samples differs slightly but not as much as for products obtained with CP14a. A slight shift for the peak maximum between individual samples can be observed.

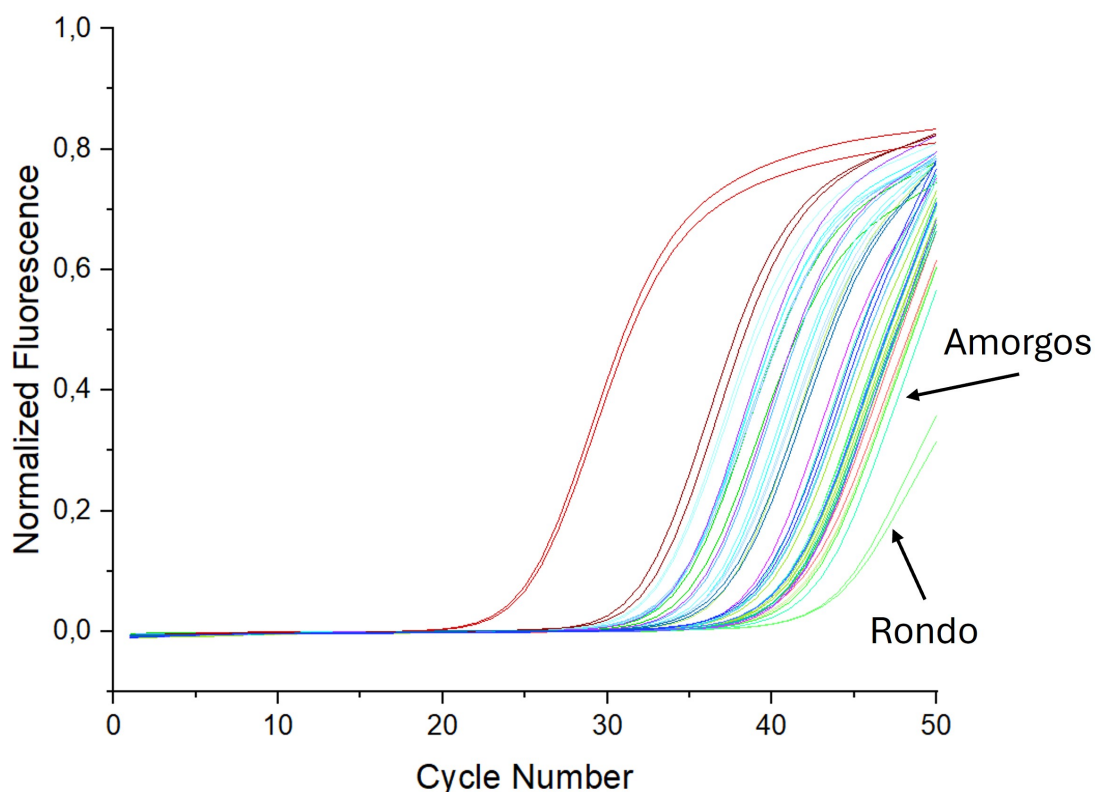


Figure 39: Amplification curves of all 25 chickpea samples analyzed in duplicates using primer pair CP14a. Identical samples share the same color. The chickpea samples with the highest CT-value Rondo (CPS6) and Amorgos (CPS16) are marked with arrows.

As the melting curves of primer pair CP14a were more complex than those generated with primer pair CP14b and showed higher potential for cultivar differentiation, additional effort was put into improving the repeatability of the assay with primer pair CP14a. A PCR run with varying magnesium chloride (1 & 2 mM) and primer concentrations (0.4, 0.6 & 0.8 μ M) was used to test if primer binding conditions could be improved. However, the results showed that this did not lead to a significant CT-value reduction, and thus, no increased magnesium chloride concentrations were used for further experiments involving CP14a. Additionally, a reduction of the annealing temperature from 57.4 to 55.4 $^{\circ}$ C was tested, as the melting temperature of the reverse primer (58.4 $^{\circ}$ C) might be too close to the annealing temperature. Despite all efforts, the repeatability did not improve. Nonetheless, primer pair CP14a could be used for cultivar differentiation as described in section 5.1.2

To gain more insight into the PCR results and to decipher the amplified product, which is no longer present in the assay using primer pair CP14b a gel electrophoresis was performed according to the description in section 4.5 (Figure 40).

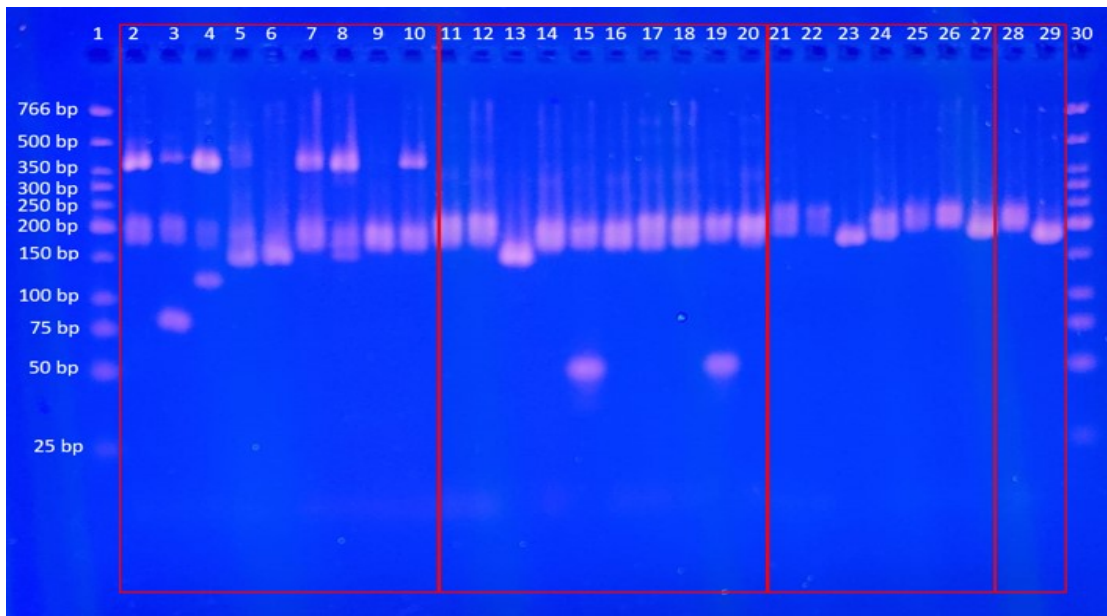


Figure 40: Picture of an agarose gel used to separate DNA bands of different sizes. Numbers 1 to 30 indicate the different lanes. The first and last lane contain the DNA ladder with the size of the individual bands indicated on the left side. The red boxes encircle samples, which were taken from the same PCR assay. For further information regarding the content of the lanes view Table [7](#)

Lanes two to ten include samples from the first test run with primer pair CP14. CP14 seemed promising for cultivar differentiation, but due to poor repeatability and the presence of two melt peaks as shown in Figure [35](#) it was assumed, that a second product might be amplified. The expected amplicon length for CP14 was estimated to be set around 156 bp. Lanes two to ten show the presence of multiple DNA bands at around 150 bp, which is the expected product. However, lanes two, three, four, five, seven, eight and ten include a second DNA band of approximately 350 bp and lanes three and four even include a third band below the size of 150 bp. Additionally, lanes six and nine, which only contained the expected product, were one of the few samples whose products showed stable repeatability throughout all tests involving primer pair CP14. This indicates that primer pair CP14 is not binding specifically.

Lanes 11 to 20 include samples amplified with primer pair CP14b. Primer pair 14b binds more specifically than primer pair CP14, but there are still some different products visible in lane 15 and lane 19.

Lanes 21 to 27 show samples amplified with primer pair CP14a. Here, no other products are observed on the gel which indicates good specificity. Lanes 28 and 29 also include samples amplified with primer pair CP14a but a lower annealing temperature was used.

They were loaded to the gel to check whether the lower temperature affected the DNA amplification in some way which could not be observed in the melting curves. However, the DNA bands show the same size range and intensity as the CP14a assay using a higher annealing temperature.

5 Results & discussion

Table 7: List of every sample loaded onto the gel depicted in Figure 40. The PCR conditions include the used primer pair, as well as the annealing temperature (T_A). Elongation temperature was uniformly set to 72 °C. A primer concentration of 0.4 μ M was used for every listed primer pair.

Lane	Sample	PCR condition	PCR date
1	Marker	-	-
2	CPS1	CP14 Primer $T_A = 51.3^\circ\text{C}$	19.01.2024
3	CPS3	CP14 Primer $T_A = 51.3^\circ\text{C}$	19.01.2024
4	CPS6	CP14 Primer $T_A = 51.3^\circ\text{C}$	19.01.2024
5	CPS11	CP14 Primer $T_A = 51.3^\circ\text{C}$	19.01.2024
6	CPS13	CP14 Primer $T_A = 51.3^\circ\text{C}$	19.01.2024
7	CPS16	CP14 Primer $T_A = 51.3^\circ\text{C}$	19.01.2024
8	CPS16	CP14 Primer $T_A = 51.3^\circ\text{C}$	19.01.2024
9	CPS18	CP14 Primer $T_A = 51.3^\circ\text{C}$	19.01.2024
10	CPS19	CP14 Primer $T_A = 51.3^\circ\text{C}$	19.01.2024
11	CPS1	CP14b Primer $T_A = 61.0^\circ\text{C}$	26.02.2024
12	CPS6	CP14b Primer $T_A = 61.0^\circ\text{C}$	26.02.2024
13	CPS13	CP14b Primer $T_A = 61.0^\circ\text{C}$	26.02.2024
14	CPS15	CP14b Primer $T_A = 61.0^\circ\text{C}$	26.02.2024
15	CPS15	CP14b Primer $T_A = 61.0^\circ\text{C}$	26.02.2024
16	CPS16	CP14b Primer $T_A = 61.0^\circ\text{C}$	26.02.2024
17	CPS20	CP14b Primer $T_A = 61.0^\circ\text{C}$	26.02.2024
18	CPS20	CP14b Primer $T_A = 61.0^\circ\text{C}$	26.02.2024
19	CPS25	CP14b Primer $T_A = 61.0^\circ\text{C}$	26.02.2024
20	CPS25	CP14b Primer $T_A = 61.0^\circ\text{C}$	26.02.2024
21	CPS1	CP14a Primer $T_A = 57.4^\circ\text{C}$	27.02.2024
22	CPS6	CP14a Primer $T_A = 57.4^\circ\text{C}$	27.02.2024
23	CPS13	CP14a Primer $T_A = 57.4^\circ\text{C}$	27.02.2024
24	CPS16	CP14a Primer $T_A = 57.4^\circ\text{C}$	27.02.2024
25	CPS16	CP14a Primer $T_A = 57.4^\circ\text{C}$	27.02.2024
26	CPS19	CP14a Primer $T_A = 57.4^\circ\text{C}$	27.02.2024
27	CPS19	CP14a Primer $T_A = 57.4^\circ\text{C}$	27.02.2024
28	CPS16	CP14a Primer $T_A = 55.4^\circ\text{C}$	01.03.2024
29	CPS16	CP14a Primer $T_A = 55.4^\circ\text{C}$	01.03.2024
30	Marker	-	-

CP14a was also tested with the second DNA extracts of all samples, but the repeatability did not improve. While the melting curves of the products of some cultivars showed continuously good repeatability through all tests with CP14a, such as CPS13, others were scattered widely, such as the melting curves of the products of CPS3, CPS4 and CPS5 (Figure 41).

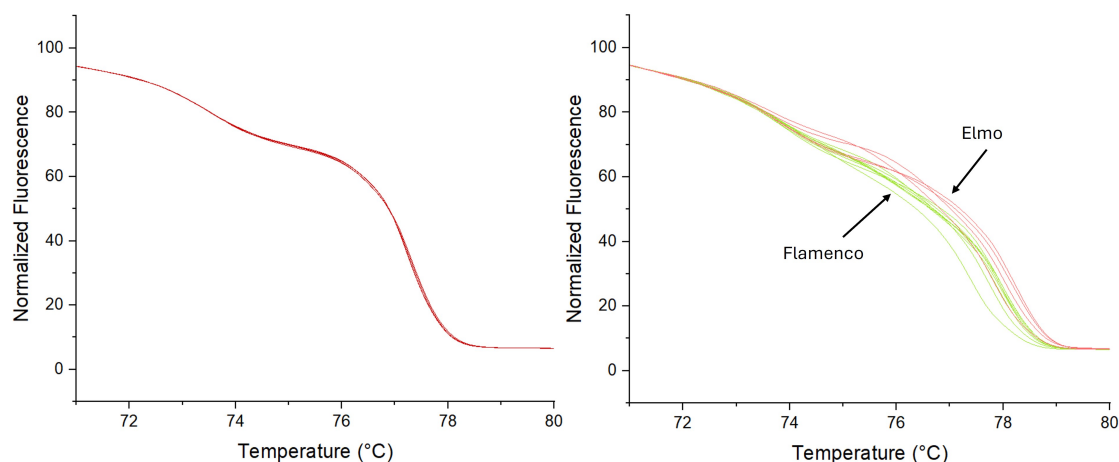


Figure 41: On the left, the melting curves of the products of CPS13 (first and second DNA extracts) with primer pair CPS14a from two separate PCR runs are shown. The right side shows the melting curves of the products of samples CPS3 and CPS4 as well as CPS5 (first and second DNA extracts) with primer pair CPS14a from two separate PCR runs. Identical samples share the same color. The melting curves scattered even though the experimental conditions were identical.

CP16 was not tested for its differentiation ability, as it was excluded during its QuantStudio™ test phase. The melt peaks generated with primer pair CP16 seemed very random and were not uniform at all as seen in Figure 42. The amplified products from the QuantStudio™ run were analyzed with a gel-electrophoresis, which showed that the primer pair either did not bind at all or bound unspecifically.

The melting curves generated with primer pair CP20 showed good properties for differentiation, but no amplification curves were recorded for all samples. The experiment was repeated on three separate days using the same conditions twice (0.4 μ M primer concentration and annealing temperature of 58.2°C) and for the third time with an increased primer concentration of 0.6 μ M. Additionally, the second DNA extract of all samples was tested with primer pair CP20. Nonetheless, no amplification curves were recorded. The melting curves resulting from the individual runs were comparable with each other, even when using the second sample extract (Figure 43). Technical difficulties or wrong Rotor-Gene Q settings can be excluded, as different primer pairs were tested in between the repetition of primer pair CP20 and no similar issues occurred.

5 Results & discussion

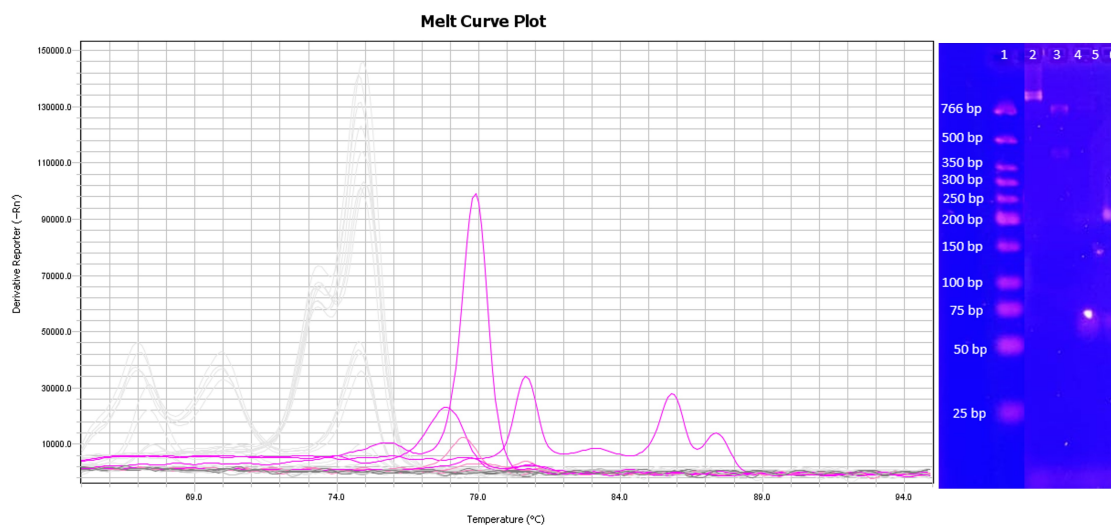


Figure 42: On the left a screenshot of the results of primer pair CP16 in the QuantStudio™ Design & Analysis software is depicted. Results are not repeatable. The right side shows the results of the conducted gel-electrophoresis. Lane one contains the DNA ladder and lanes two to six contain chickpea DNA amplified with CP16.

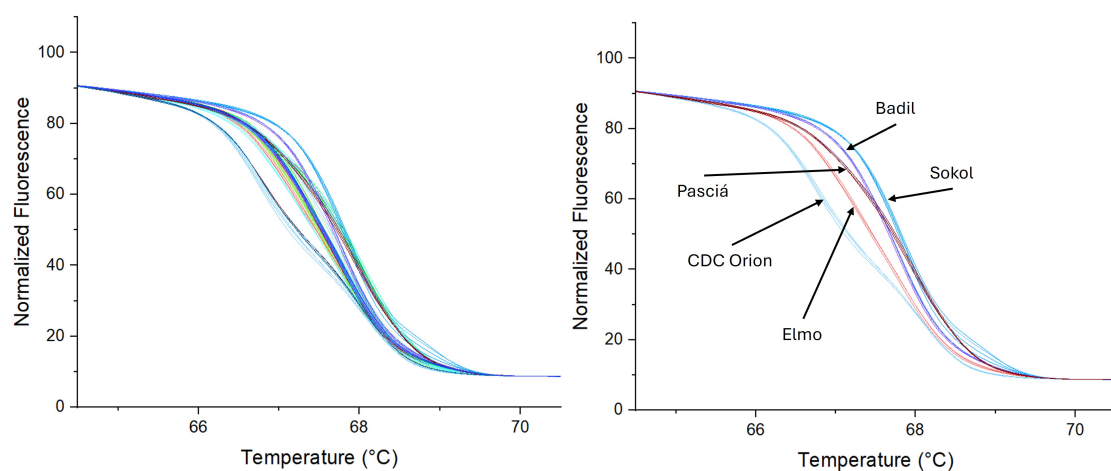


Figure 43: On the left, all melting curves of the products of all samples with primer pair CP20 are depicted. Two different DNA extracts analyzed with two different PCR runs were compared. For each sample four melting curves are shown as the first and the second extract were analyzed in duplicates. Identical samples share the same color. On the right, a few samples whose amplicons produced different melting curve shapes were selected from the graphic on the left to showcase the good repeatability between the two extracts and different PCR runs.

Another gel electrophoresis was conducted to check if the melting curves generated with primer pair CP20 were produced by the amplified target region. The expected amplicon length for primer pair CP20 was estimated to be around 198 bp. DNA bands of various sizes between 200 and 300 bp were visible on the gel indicating that the amplification of the target regions was successful (Figure 44). While some other DNA bands of larger size are visible on the gel, their concentration is rather low compared to the target region, whose DNA bands show a high intensity. Overall, the gel provided more insight into the PCR results of primer pair CP20. While it is still unclear why the fluorescence signal did not increase during amplification, the generated melting curves are considered to be reliable. Nonetheless, primer pair CP20 was not used for the final result due to said problems during amplification.

The middle section of the gel (lanes 9 - 22) refers to an analysis of primer pair SNP7, which will be further explained in section 5.2.1.

The right side of the gel (lanes 23 - 29) tested individual samples analyzed with primer pair CP8. This served as a test whether differences in amplicon length could be observed. Products of individual samples with high differences in their melting curve and, thus, in their microsatellite length were handpicked for this purpose. However, no real size differences could be observed on the gel, as all bands were set at around 200 bp.

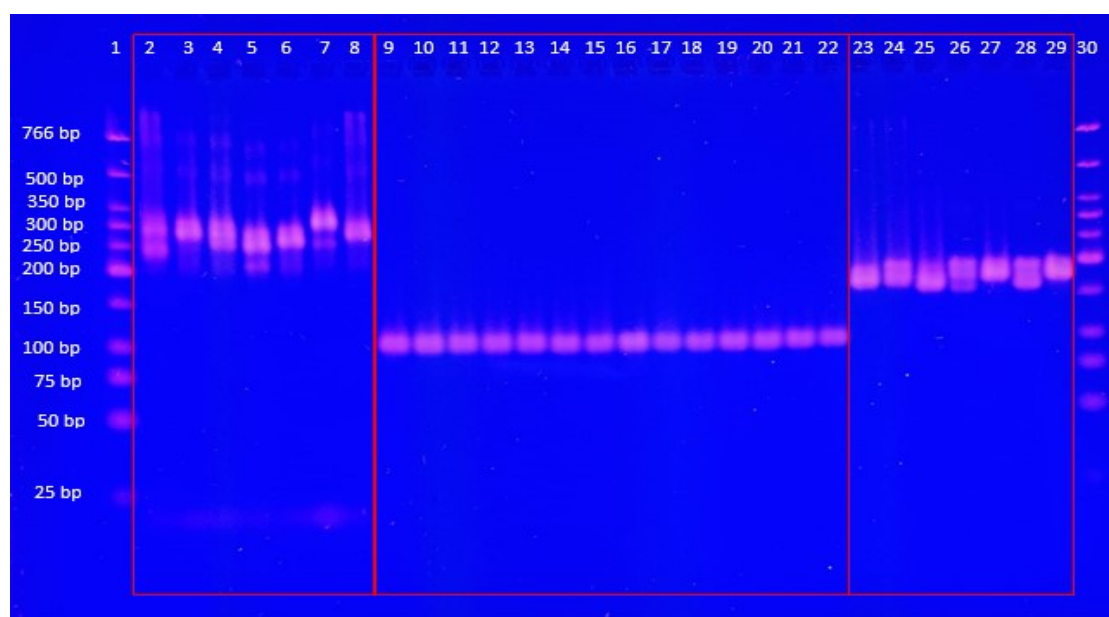


Figure 44: Picture of an agarose gel used to separate DNA bands of different sizes. Numbers 1 to 30 indicate the different lanes. The first and last lane contain the DNA ladder with the size of the individual bands indicated on the left side. The red boxes encircle samples, which were taken from the same PCR assay. For further information regarding the content of the lanes view Table 8.

5 Results & discussion

Table 8: List of every sample loaded onto the gel depicted in Figure 44. The PCR conditions include the used primer pair, as well as the primer concentration, annealing temperature (T_A) and elongation temperature (T_E).

Lane	Sample	PCR condition	PCR date
1	M 1/4	-	-
2	CPS9	CP20 Primer = 0.4 μ M T_A = 58.2 °C T_E = 68 °C	31.01.2024
3	CPS12	CP20 Primer = 0.4 μ M T_A = 58.2 °C T_E = 68 °C	31.01.2024
4	CPS17	CP20 Primer = 0.4 μ M T_A = 58.2 °C T_E = 68 °C	31.01.2024
5	CPS20	CP20 Primer = 0.4 μ M T_A = 58.2 °C T_E = 68 °C	31.01.2024
6	CPS22	CP20 Primer = 0.4 μ M T_A = 58.2 °C T_E = 68 °C	31.01.2024
7	CPS24	CP20 Primer = 0.4 μ M T_A = 58.2 °C T_E = 68 °C	31.01.2024
8	CPS25	CP20 Primer = 0.4 μ M T_A = 58.2 °C T_E = 68 °C	31.01.2024
9	CPS2	SNP7 Primer = 0.2 μ M T_A = 53.3 °C T_E = 72 °C	21.02.2024
10	CPS2	SNP7 Primer = 0.2 μ M T_A = 53.3 °C T_E = 72 °C	21.02.2024
11	CPS8	SNP7 Primer = 0.2 μ M T_A = 53.3 °C T_E = 72 °C	21.02.2024
12	CPS8	SNP7 Primer = 0.2 μ M T_A = 53.3 °C T_E = 72 °C	21.02.2024
13	CPS11	SNP7 Primer = 0.2 μ M T_A = 53.3 °C T_E = 72 °C	21.02.2024
14	CPS11	SNP7 Primer = 0.2 μ M T_A = 53.3 °C T_E = 72 °C	21.02.2024
15	CPS13	SNP7 Primer = 0.2 μ M T_A = 53.3 °C T_E = 72 °C	21.02.2024
16	CPS13	SNP7 Primer = 0.2 μ M T_A = 53.3 °C T_E = 72 °C	21.02.2024
17	CPS18	SNP7 Primer = 0.2 μ M T_A = 53.3 °C T_E = 72 °C	21.02.2024
18	CPS18	SNP7 Primer = 0.2 μ M T_A = 53.3 °C T_E = 72 °C	21.02.2024
19	CPS21	SNP7 Primer = 0.2 μ M T_A = 53.3 °C T_E = 72 °C	21.02.2024
20	CPS21	SNP7 Primer = 0.2 μ M T_A = 53.3 °C T_E = 72 °C	21.02.2024
21	CPS22	SNP7 Primer = 0.2 μ M T_A = 53.3 °C T_E = 72 °C	21.02.2024
22	CPS25	SNP7 Primer = 0.2 μ M T_A = 53.3 °C T_E = 72 °C	21.02.2024
23	CPS5	CP8 Primer = 0.4 μ M T_A = 53.3 °C T_E = 68 °C	05.12.2023
24	CPS8	CP8 Primer = 0.4 μ M T_A = 53.3 °C T_E = 68 °C	05.12.2023
25	CPS11	CP8 Primer = 0.4 μ M T_A = 53.3 °C T_E = 68 °C	05.12.2023
26	CPS17	CP8 Primer = 0.4 μ M T_A = 53.3 °C T_E = 68 °C	05.12.2023
27	CPS21	CP8 Primer = 0.4 μ M T_A = 53.3 °C T_E = 68 °C	05.12.2023
28	CPS23	CP8 Primer = 0.4 μ M T_A = 53.3 °C T_E = 68 °C	05.12.2023
29	CPS24	CP8 Primer = 0.4 μ M T_A = 53.3 °C T_E = 68 °C	05.12.2023
30	M 1/4	-	-

At first, primer pairs with microsatellites composed of only A and T (CP5, CP6, CP7, CP8, CP10, CP11, CP15, CP17, CP18, CP19, and CP20) were problematic, as the melting curves were scattering, and the amplification did not work properly. The results improved when the elongation temperature was adjusted from 72 °C to 68 °C. Additionally, the primer concentration was reduced from 0.6 μ M to 0.2 μ M, to minimize the risk of non-specific primer binding (Figure 45).

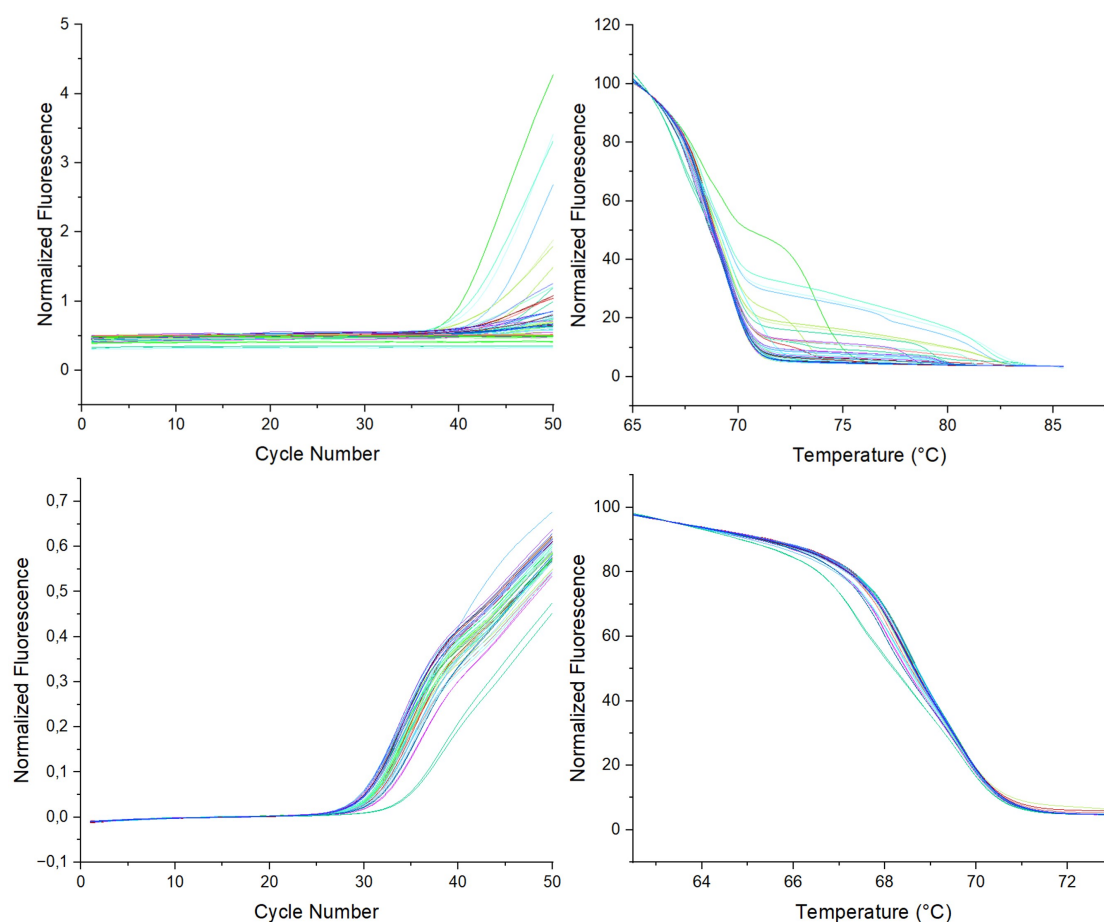


Figure 45: The top row shows amplification curves (left) and melting curves (right) generated from the products of all chickpea samples analyzed in duplicates with primer pair CP7. A primer concentration of $0.6\text{ }\mu\text{M}$ and an annealing temperature of $72\text{ }^{\circ}\text{C}$ were used. The bottom row shows the amplification curves (left) and melting curves (right) with primer pair CP7 and optimized conditions. The primer concentration and annealing temperature were reduced to $0.2\text{ }\mu\text{M}$ and $68\text{ }^{\circ}\text{C}$ respectively. Identical samples share the same color.

5.1.2 Chickpea cultivar differentiation based on microsatellites

The selected primer pairs were tested for their ability to differentiate the 24 analyzed cultivars in this thesis. The evaluation of said ability was briefly explained in section [4.8](#). The melting curves of each cultivar were compared with one another, and it was decided whether the cultivars could not be differentiated (red), could probably be differentiated (yellow) or could reliably be differentiated (green). As an overview, an excel data-sheet is created for each primer pair, where the colors red, yellow or green are used to assign each sample pair a property (differentiated - not differentiated).

Most of the obtained results are based on the analysis of one singular DNA extract. To account for possible deviations during DNA extraction a second DNA extract of all samples was prepared and tested. The second extracts were tested with the best

microsatellite primer pair CP8 and served to sort out yellow-marked samples and assign them to either red or green. When comparing the melting curves of the second extracts with the first ones using primer pair CP8, a slight shift of the melting curves to the left could be observed for all samples (Figure 46). It was assumed that technical fluctuations during the PCR led to the shift as both extracts were analyzed in two separate runs. This was further confirmed by a later performed PCR run, which involved the analysis of the first and the second DNA extracts in the same run with primer pair CP8. For the resulting melting curves no prior noticed shift could be observed (Figures 47, 48).

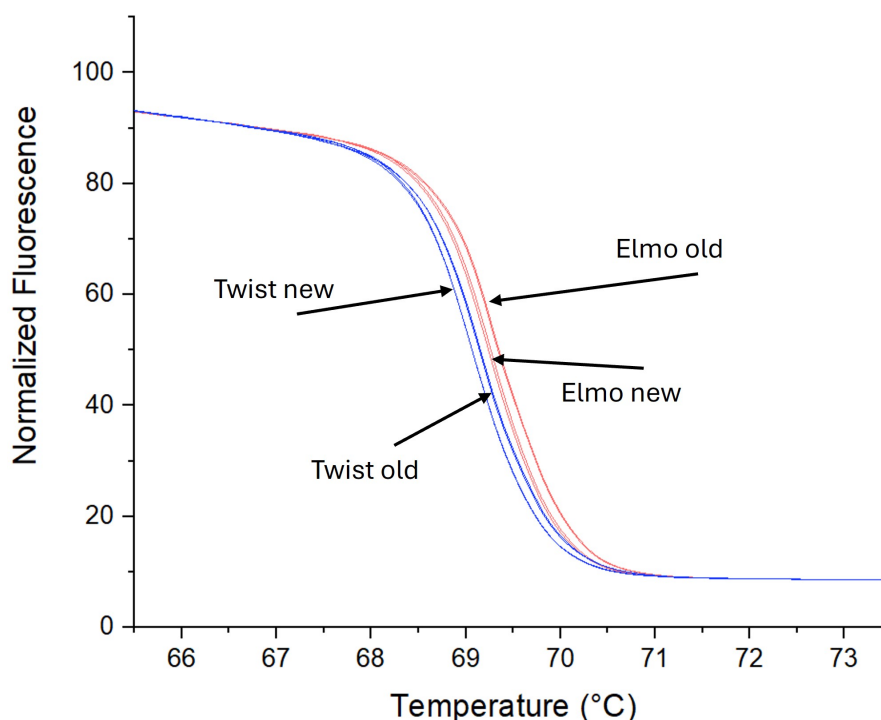


Figure 46: Melting curves of the products of cultivars Elmo and Twist amplified with primer pair CP8. Two DNA extracts were tested in two separate PCR runs and their melting curves were combined in this figure. A slight shift to the left can be observed for the second DNA extracts.

To further elaborate, which samples are differentiable, and which are not, and in order to be able to assign the yellow fields to either red (not differentiated) or green (differentiated) a temperature difference of 0.1 °C was chosen as the lower limit to mark samples as differentiated (Figures 47, 48). This limit was elaborated with a PCR, which involved the analysis of the first and the second DNA extracts in the same run with primer pair CP8 to account for possible scattering due to differences during sample preparation, DNA extraction, and technical fluctuations by the thermocycler. All primer pair excel sheets were reevaluated to match the newly assigned criterion. Thus, some samples were marked red (not differentiable) even though their melting curves could be distinguished to account for possible scattering due to experimental differences (different DNA extracts

or extraction methods, different thermocycler) and other external factors such as different scientists performing the experiment.

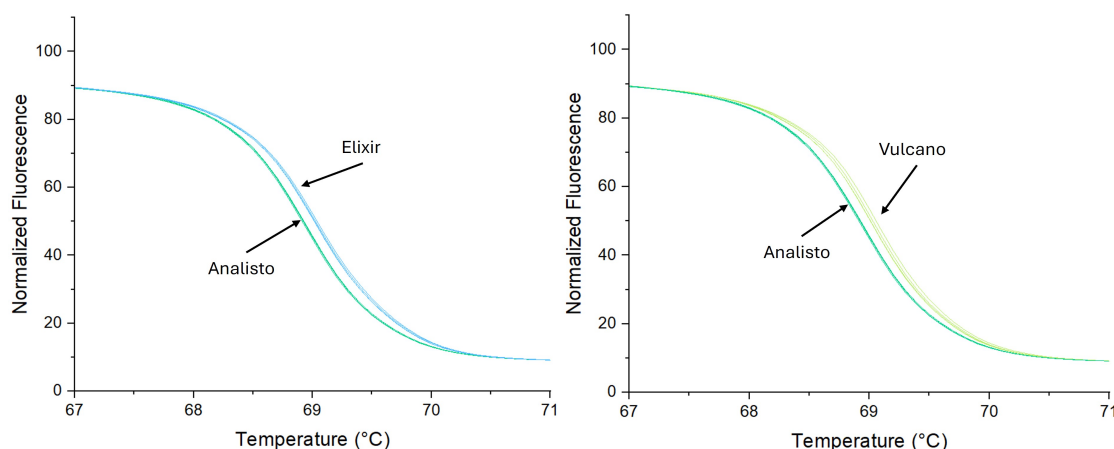


Figure 47: Representation of the melting curves of the products of two sample pairs with the first and the second DNA extracts which would be marked as green (differentiable) by the newly set criterion. Their melting curves differ by 0.15°C . Identical samples share the same color. Samples were analyzed with primer pair CP8 and optimized settings as listed in Table 6.

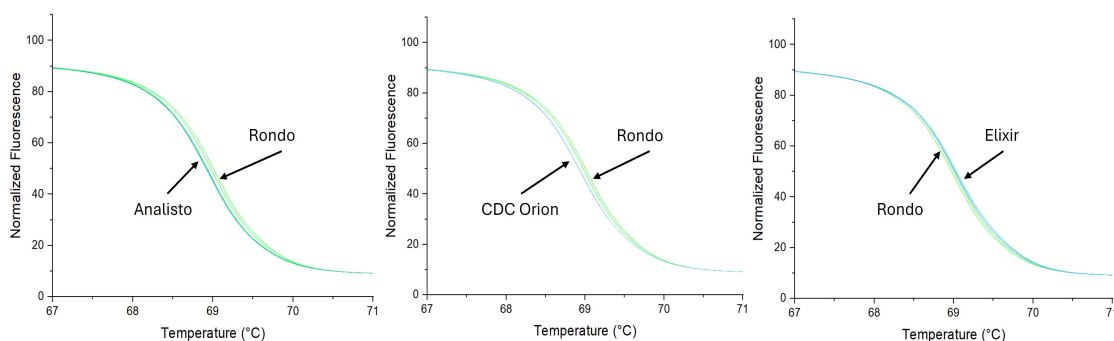


Figure 48: Representation of the melting curves of the products of three sample pairs with the first and the second DNA extracts which would be marked as red (not differentiable) by the newly set criterion. Their melting curves differ by less than 0.1°C . Identical samples share the same color. Samples were analyzed with primer pair CP8 and optimized settings as listed in Table 6.

Since the analyzed chickpea extracts were not evenly diluted to a certain DNA concentration but rather diluted 1:100 regardless of their initial concentration, the same PCR run was used to compare the melting curves of the first and the second DNA extracts, which were diluted 1:100 and the first and the second DNA extracts, which were diluted to a concentration of $2.5\text{ ng}/\mu\text{L}$ with nuclease-free water. As seen in Figure 49, the melting curves of the first and the second DNA extracts were identical. Thus, the influence of the different DNA sample concentrations and, thus, sample matrix on melting behaviour can most likely be excluded.

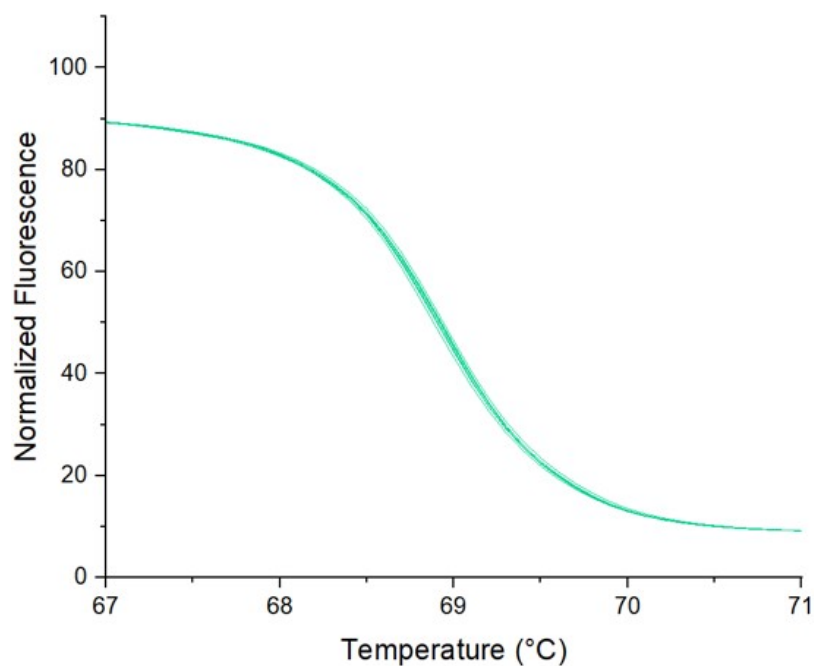


Figure 49: Melting curves of the first and the second DNA extracts of the products of chickpea cultivar Analisto with primer pair CP8. The first and the second extracts were diluted 1:100 or to 2.5 ng/ μ L. Regardless of their initial DNA concentration, the melting curves overlap.

The excel sheets of individual primer pairs were compared to each other, and complementary primer pairs were grouped together to create a final excel sheet for microsatellites by using as few primer pairs as possible to differentiate most cultivars (Figure 50). All but twelve cultivar pairs could be differentiated with the tested primer pairs. Some of the not differentiable pairs could be resolved with SNPs as listed in Figure 61.

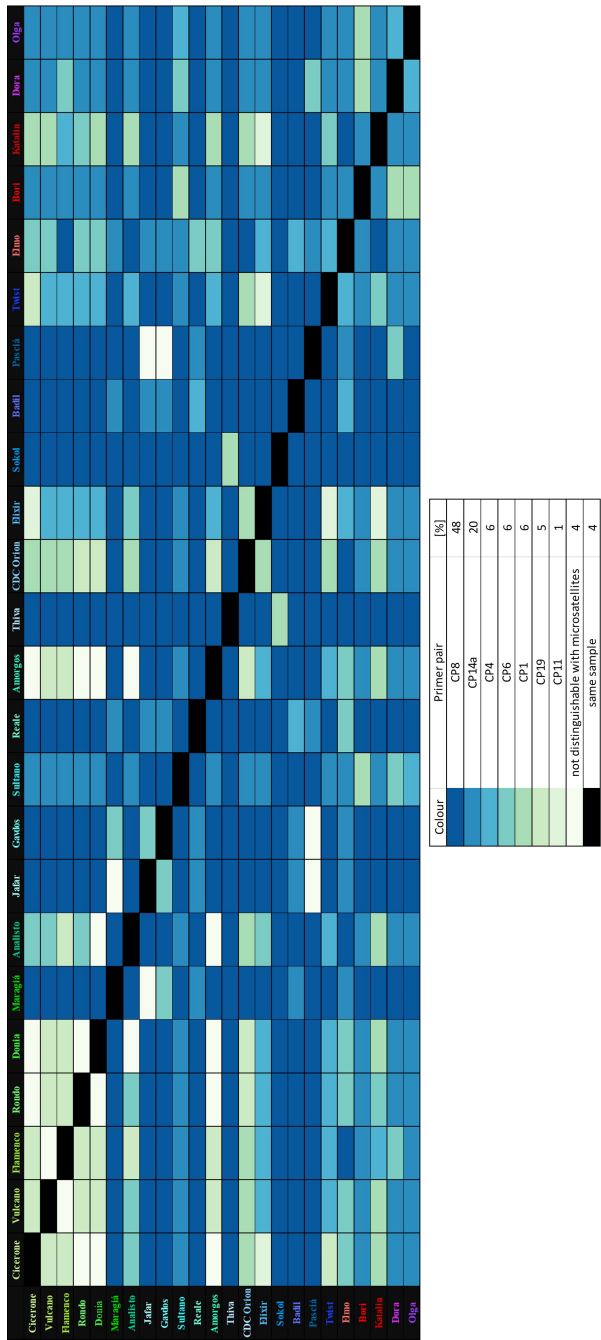


Figure 50: By comparing the excel sheets of different primer pairs and grouping complementary ones a final excel sheet for microsatellites is created. Samples are sorted by their type, starting with kabuli, followed by desi and gulabi. All but twelve cultivar pairs could be differentiated using the listed primer pairs in the legend.

5 Results & discussion

The following pages show examples of melting curves generated with primer pairs used in the final excel data sheet (Figures 51 - 57).

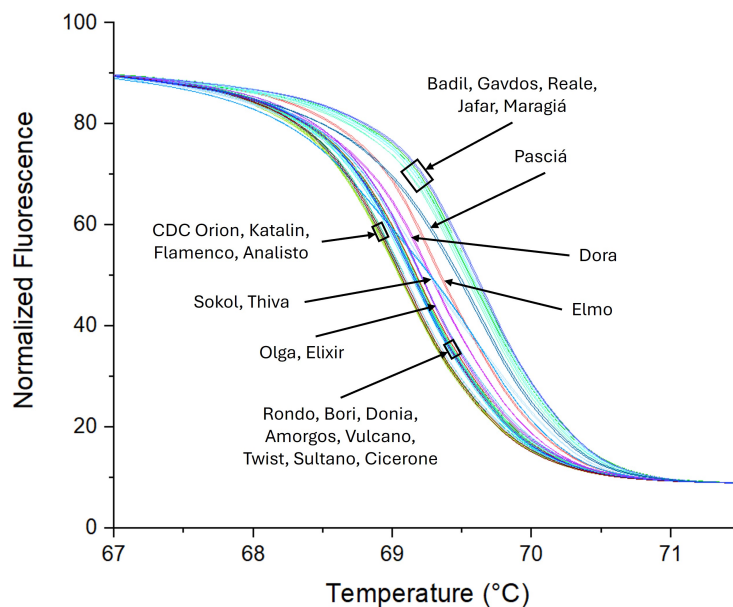


Figure 51: Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair CP8. Identical samples share the same color.

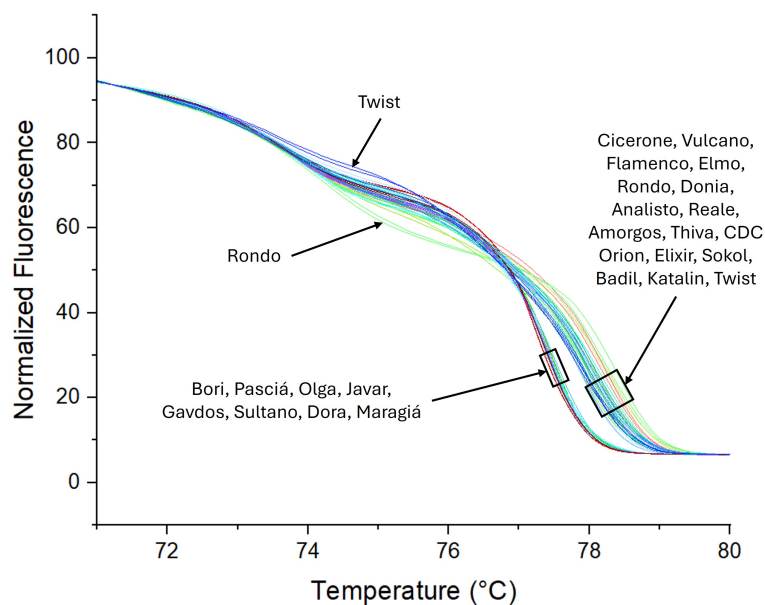


Figure 52: Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair CP14a. Identical samples share the same color.

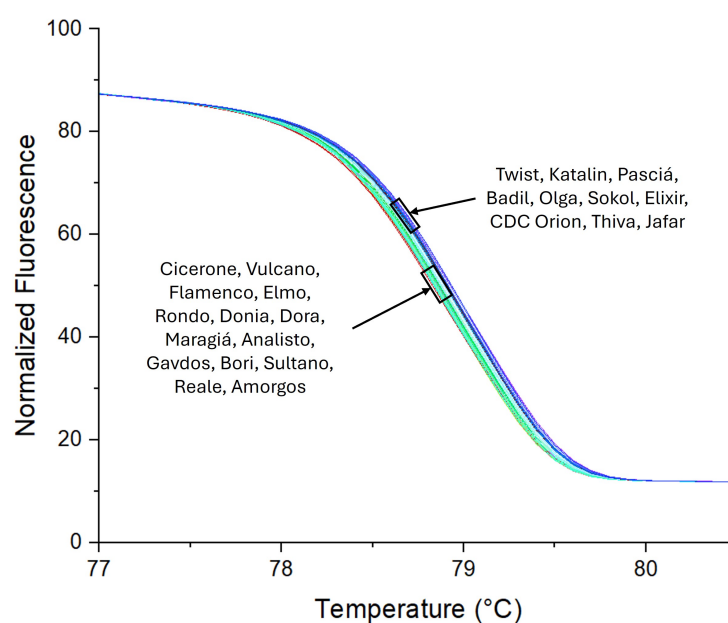


Figure 53: Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair CP4. Identical samples share the same color.

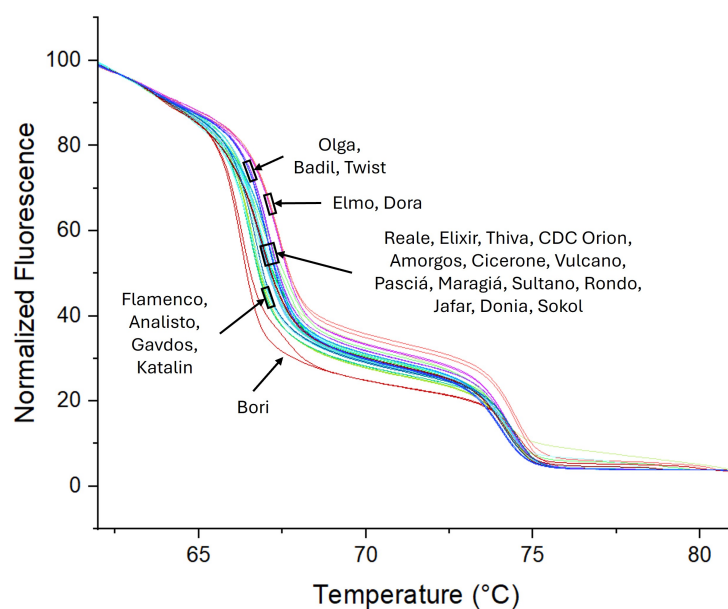


Figure 54: Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair CP6. Identical samples share the same color.

5 Results & discussion

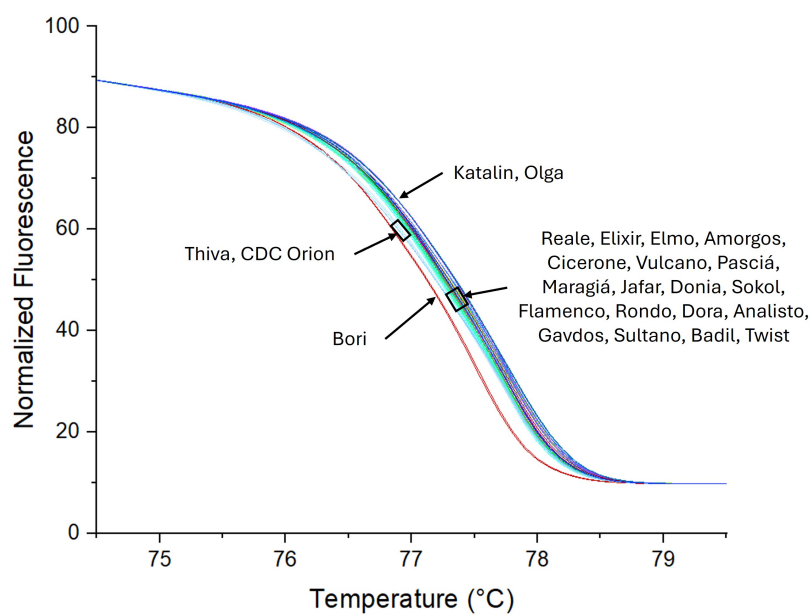


Figure 55: Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair CP1. Identical samples share the same color.

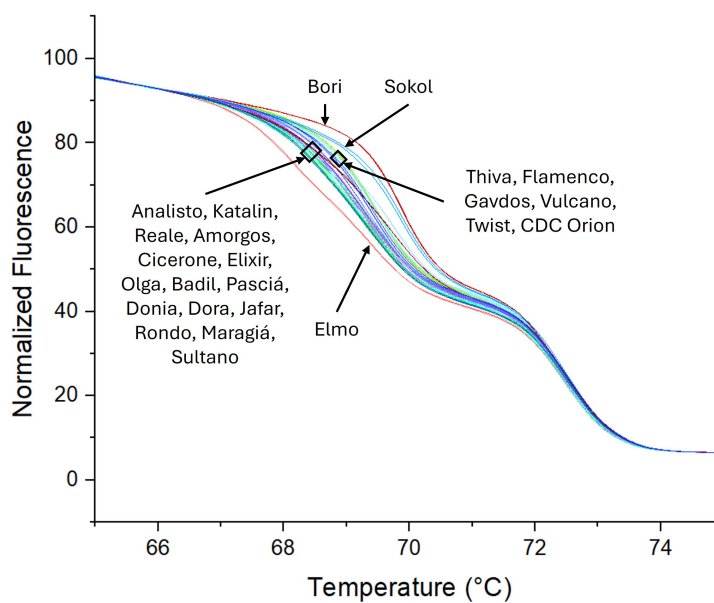


Figure 56: Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair CP19. Identical samples share the same color.

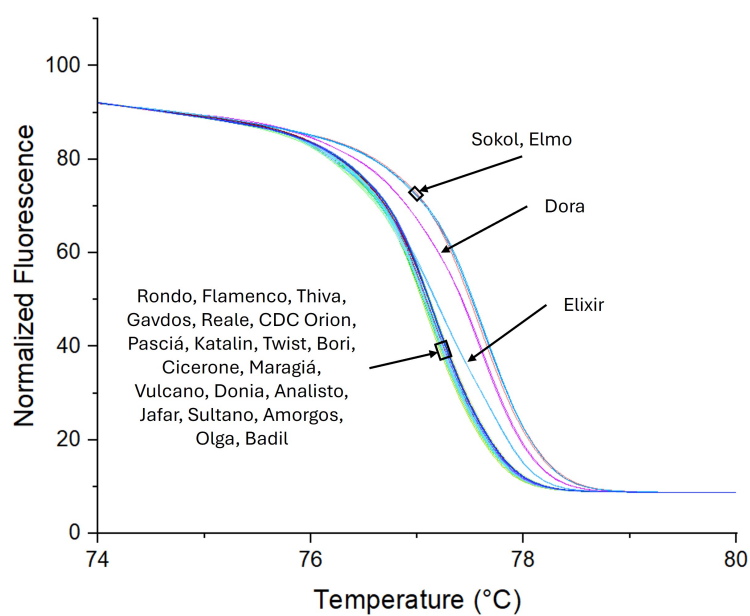


Figure 57: Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair CP11. Identical samples share the same color.

5.2 Single nucleotide polymorphism

5.2.1 Suitability of primer pairs

Primer pairs targeting SNPs were newly designed as explained in section 4.2.3. Similar to microsatellites, primer pairs were tested *in silico* and filtered for specific criteria as listed in sections 3.1 and 4.2. A summary of the *in silico* test results of primer pair SNP2, which was the most promising regarding its cultivar differentiation ability, is listed in Figure 58.

SNP2	Primer sequence	Length (bp)	GC content	TM (°C)
Fw	CTGGTCCTTGAGATCACTGT	20	50	58.4
Rev	GTTCAAGGGCAGCTGTATT	20	50	58.4
Sec				
Dim	Forward-Dimers Delta G: -4.62 kcal/mol Base Pairs: 4 5' CTGGTCCTTGAGATCACTGT 3' GTTCAAGGGCAGCTGTATT Reverse-Dimers Delta G: -10.24 kcal/mol Base Pairs: 6 5' GTTCAAGGGCAGCTGTATT 3' CTGGTCCTTGAGATCACTGT Forward + Reverse Dimers Delta G: -4.67 kcal/mol Base Pairs: 3 5' CTGGTCCTTGAGATCACTGT 3' GTTCAAGGGCAGCTGTATT Amplicon length: 84 bp			

CTGGTCCTTGAGATCACTGTACCTTGCATCCTCAATTTTCCTAATCTCGTCCCTTGAACAATACAGCTGCCCTCTGAAC

Figure 58: Overview of the *in silico* test results of primer pair SNP2 including primer sequence, length, GC content, melting temperature (T_M), amplicon length, secondary structures (Sec) and potentially formed dimers (Dim). At the bottom, the location of the forward and reverse primer, and the individual SNPs are marked with yellow and cyan, respectively.

Nine different primer pairs were tested in the lab as described in sections 4.7 and 4.8. The sequences and used experimental conditions of the primer pairs are listed in Table 9. Primer pairs were named with the abbreviation SNP and numbers one to nine. Additionally, the table lists the potential of the singular primer pairs to differentiate the tested chickpea cultivars.

Table 9: All primer pairs tested for cultivar differentiation targeting SNPs in this thesis including their sequence, their used concentration (Conc.), annealing temperature (T_A) and whether they were suitable for differentiation (Y) or not (N). For primer pairs marked as (N) the melting curves of all cultivars were overlapping.

Name	fw-sequence	rev-sequence	Conc. (μM)	T _A (°C)	Differentiation
SNP1	ACATAGAAAGTCTCTGTTTGC	ACATATTAGCAGCCTATTCA	0.2	51.3	Y
SNP2	CTGTGCTCTTGAGATCACTGT	GTTCAGAGGGGAGCTGTATT	0.2	57.4	Y
SNP3	CTTCAAGATATCAATGCAGC	TGATGGATGAGTTTCATTAGG	0.4	53.3	Y
SNP4	TTTACAACAGTGAAGCAGCAG	ACGAGGATTGCAAGTAGTGA	0.2	55.4	Y
SNP5	TAGCATCGTCAAATGCAGCA	TCAGCTCGCATTTGCTGTCTA	0.2	55.4	N
SNP6	GAATCTACTAATGAAGGTTA	GCAGCTACAATATTATAGGT	0.4	49.2	N
SNP7	ATTGGATCTTCTCTTTTGGC	ATTGTTGACAGTTCTGGTGC	0.2	53.3	Y
SNP8	GCTCTGTGCTTCTTGAGT	CTTCTAACCGATTGCGGTGG	0.2	59.5	Y
SNP9	AGTAACTCCTCAACCTGGAG	CCATGTCCAGTGGGAAGATT	0.4	57.4	N

Similar to primer pair CP14 primer pair SNP7 seemed promising for cultivar differentiation but did not provide repeatable melting curves. All cultivars were analyzed with primer pair SNP7 twice on two separate days and the melting curves seemed to scatter randomly. Nonetheless, the formation of three distinct groups can be observed (Figure 59). Although the melting curves were scattered to some degree, the three groups can still be clearly identified, which is why no further adjustments were made.

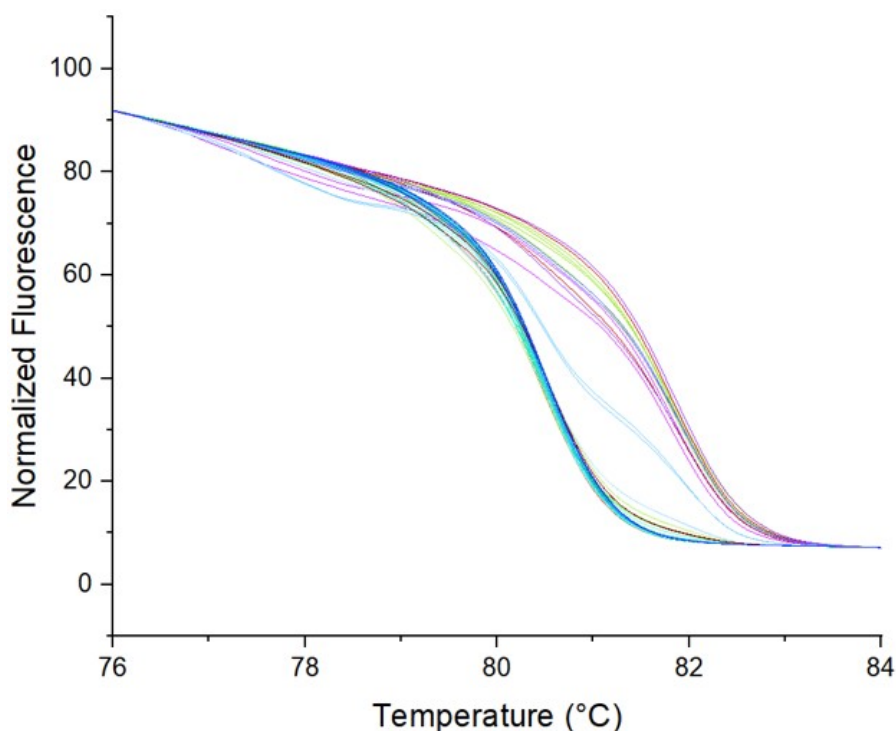


Figure 59: This figure shows the melting curves of the products of all 25 chickpea samples analyzed in duplicates using primer pair SNP7. Identical samples share the same color.

Additionally, products from some cultivars amplified with primer pair SNP7 were analyzed by a gel electrophoresis to check if any other unspecific regions are amplified, which might account for the melting curve scattering. However, on the gel only DNA bands with the expected amplicon length of 100 bp were visible (Figure 44).

5.2.2 Chickpea cultivar differentiation based on SNPs

Similar to the primer pairs targeting microsatellites the selected SNP primer pairs were tested for their ability to differentiate the 24 analyzed cultivars. The evaluation of said ability was already explained in sections 4.8 and 5.1.2. Cultivar pairs which were marked yellow were reevaluated with the second DNA extracts to assign them either red (not differentiable) or green (differentiable) colors.

Primer pair SNP2 was analyzed with the first and the second DNA extracts of all chickpea samples similar to primer pair CP8 to check for possible deviations during DNA extraction. In comparison to primer pair CP8, no shift was observed for the melting curves of the products of all samples as the melting curves of the first and the second DNA extracts were overlapping (Figure 60).

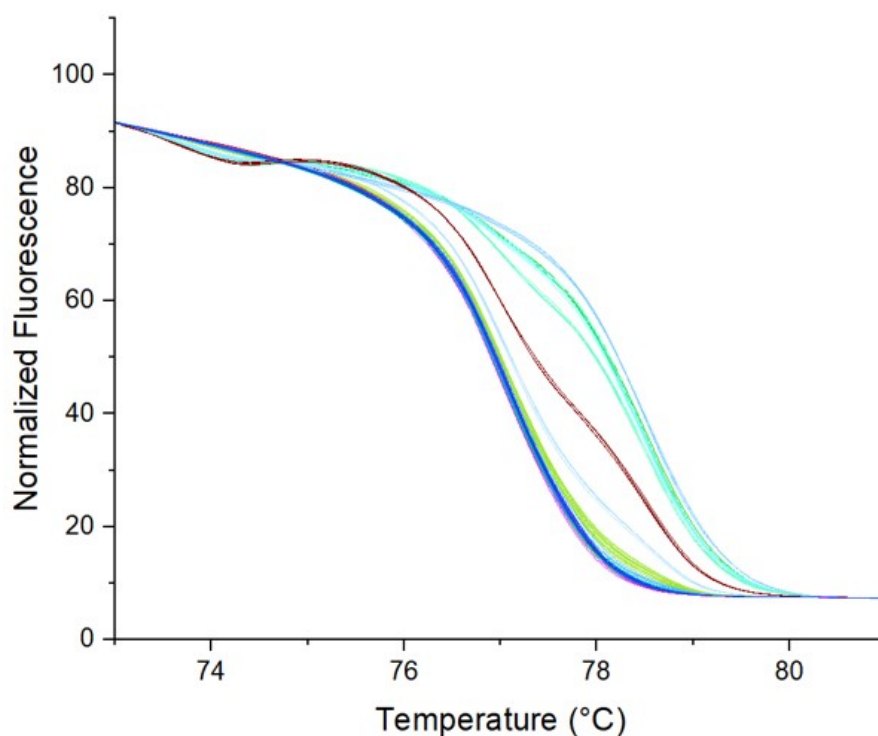


Figure 60: Melting curves of the products of all 25 chickpea samples amplified with primer pair SNP2. This figure combines the melting curves of two separate PCR runs with two different DNA extracts. The melting curves produced by the amplicons of the same sample overlap completely and no shift can be observed. Identical samples share the same color.

Primer pair SNP8 was the only primer pair, which was able to distinguish the chickpea types. A different melting behaviour was observed for the kabuli type. The only exception is the cultivar Flamenco, which shares the same melting behaviour as the gulabi and desi types (Figure 64). The sequencing outcome supports this result (Figure 73).

The excel sheets of individual primer pairs were compared to each other, and complementary primer pairs were grouped together to create a final excel sheet for SNPs by using as few primer pairs as possible to differentiate most cultivars (Figure 61). All but 19 cultivar pairs could be differentiated with the tested primer pairs.

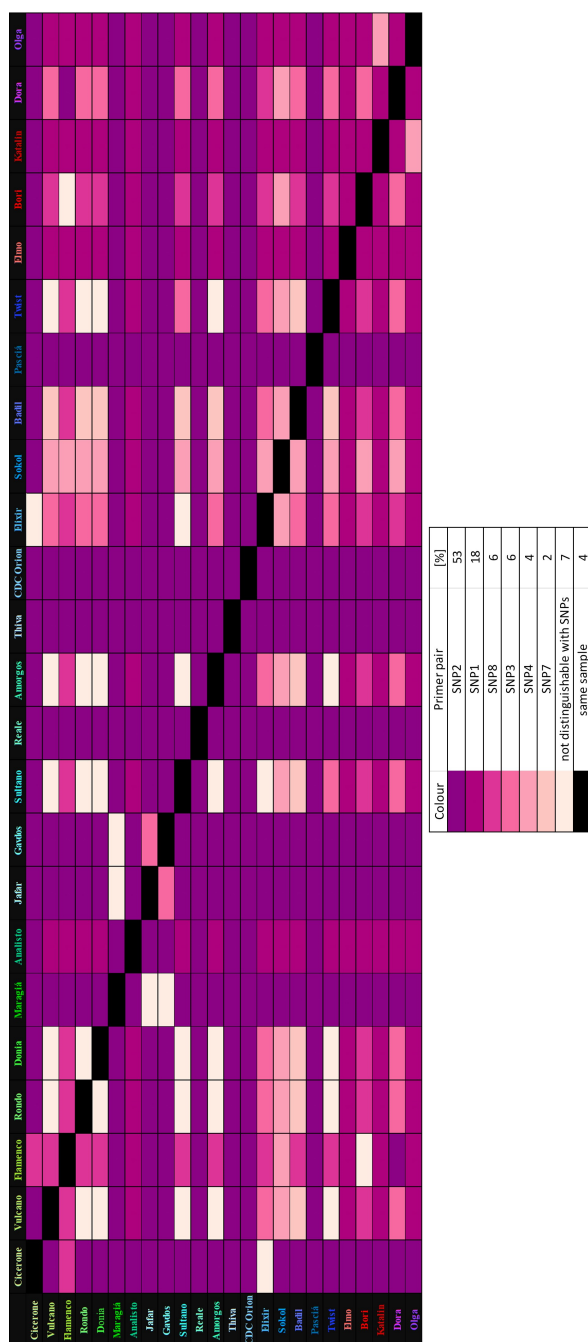


Figure 61: By comparing the excel sheets of different primer pairs and grouping complementary ones a final excel sheet for SNPs is created. Samples are sorted by their type, starting with kabuli, followed by desi and gulabi. All but 19 cultivar pairs could be differentiated using the listed primer pairs in the legend.

5.2 Single nucleotide polymorphism

The following pages show examples of melting curves generated with primer pairs used in the final excel data sheet (Figures 62 - 67).

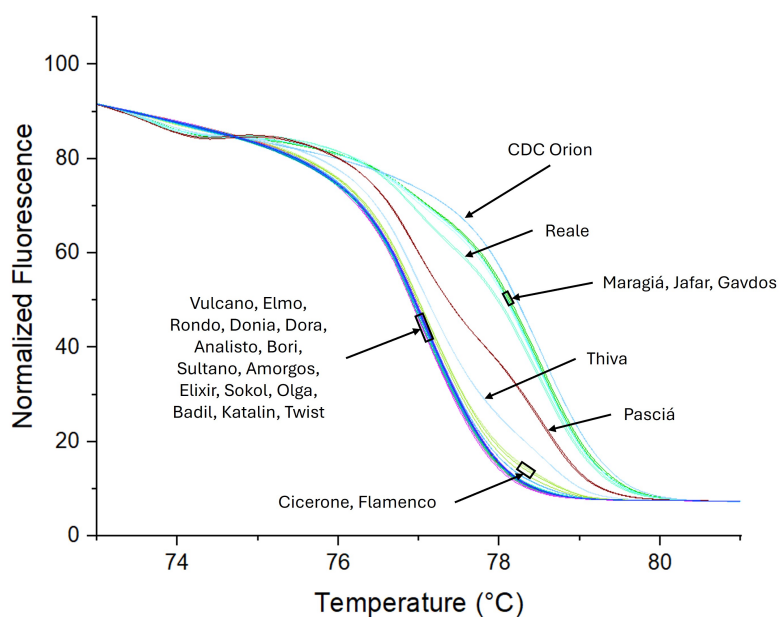


Figure 62: Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair SNP2. Identical samples share the same color.

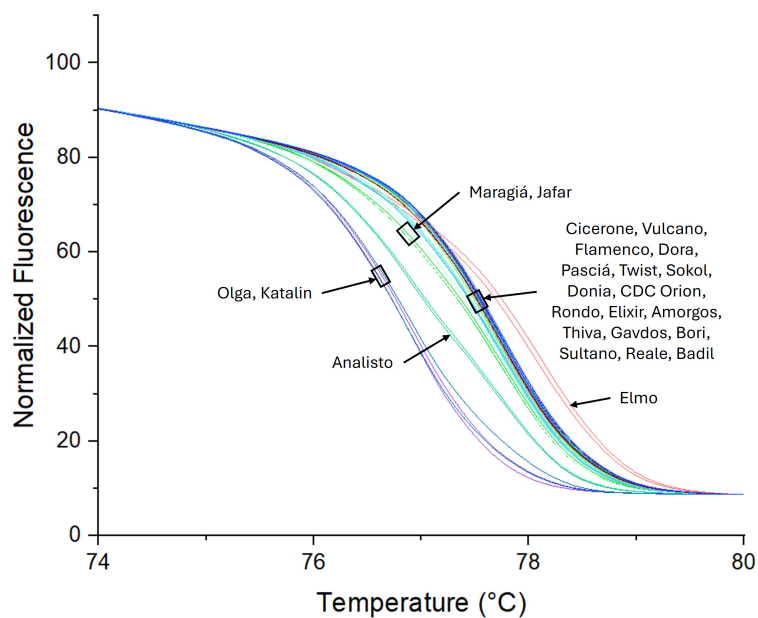


Figure 63: Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair SNP1. Identical samples share the same color.

5 Results & discussion

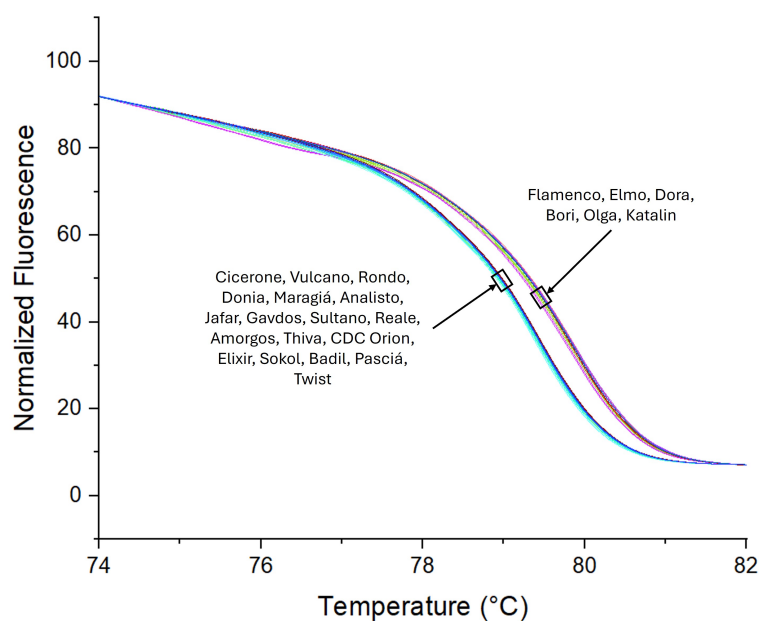


Figure 64: Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair SNP8. This is the only primer pair which could differentiate between different chickpea types. Identical samples share the same color.

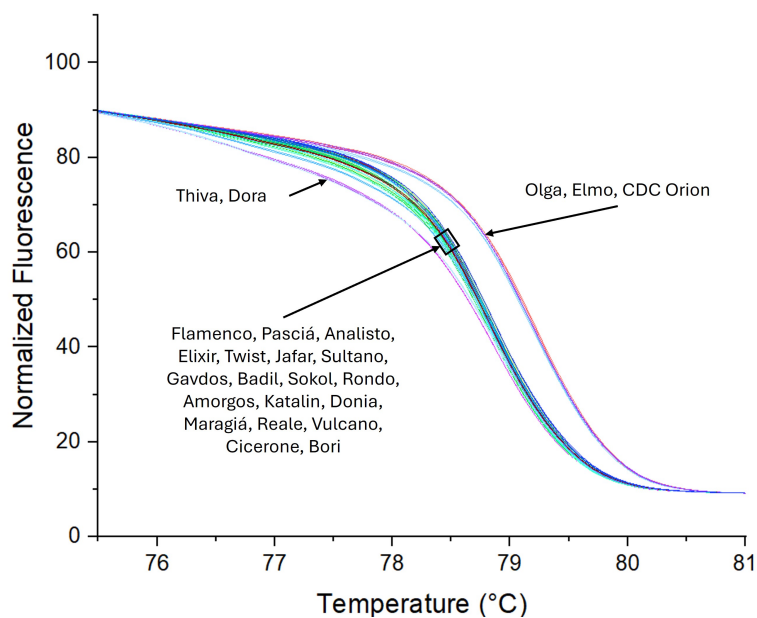


Figure 65: Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair SNP4. Identical samples share the same color.

5.2 Single nucleotide polymorphism

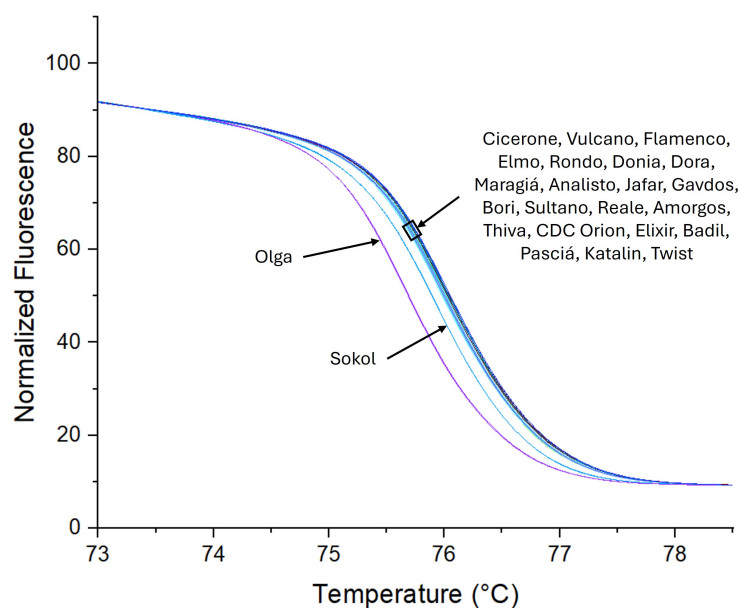


Figure 66: Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair SNP3. Identical samples share the same color.

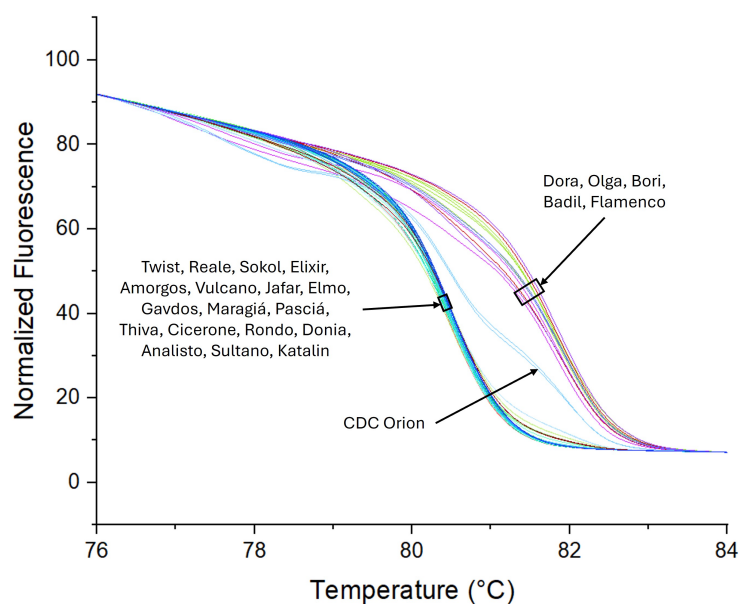


Figure 67: Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair SNP7. Identical samples share the same color.

5.2.3 Sequencing results

To gain more insight into the position and shape of the melting curves the amplicons were sequenced using the pyrosequencing technique. The basics behind this method and how the sequencing runs were performed are described in sections 3.7 and 4.9 respectively.

For interesting primer pairs with high potential for cultivar differentiation, such as SNP2 and SNP1, the PCR products of all cultivars were sequenced. For other primer pairs where the melting curves produced by the amplicons of the cultivars could be grouped, such as SNP8 or SNP3, only a few randomly picked samples from each group were sequenced. The PCR products of some cultivars amplified with SNP7 showed bad repeatability, and therefore, both duplicates of said cultivars were sequenced. For cultivars with good repeatability, only one of both duplicates was sequenced.

There are three possible outcomes for SNPs. As already listed in section 3.9 SNPs can be homozygous in two ways (either only one base or only the other base is present) or heterozygous (both possible bases are present). The melt peaks generated with primer pair SNP7 provide a good example. Since the amplicons generated with SNP7 include three SNPs and chickpeas are diploid organisms, more complex melting behaviour is observed. Differences in SNPs can be observed by a reduction or increase in the melting temperature required to separate the DNA strands which is showcased in Figure 68. The three SNPs in Elixir are homozygous with T/T, A/A and T/T and thus melt at a lower temperature compared to Dora, which is homozygous with G/G, G/G and C/C. All three SNPs in CDC Orion are heterozygous (G/T, A/G and C/T) and form two melt peaks with varying intensity.

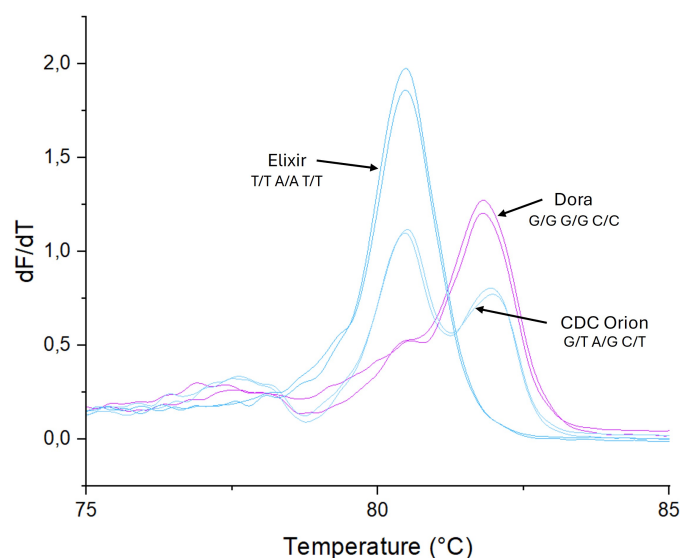


Figure 68: Melt peaks of the products of three chickpea cultivars amplified with primer pair SNP7. All three SNPs in Elixir and Dora are homozygous and show one distinct peak. Since Elixir is homozygous with A and T the melt peak is located at a lower temperature compared to Dora, whose SNPs are composed of G and C. CDC Orion is heterozygous for all three SNPs and thus two peaks of varying intensity can be observed.

The following graphs contain the sequencing results of the PCR products of individual primer pairs (Figures 69- 73). The quality of the results is color-coded as already explained in section 4.9. SNPs which are marked blue are reliable, and yellow-marked SNPs are still correct to some degree but should be checked. Red-marked SNPs are not recognized correctly and should be repeated. In this thesis, all red-marked SNPs were sequenced two or three times but remained red. Thus, red-marked results are to be treated with caution.

When comparing the sequencing results with the melting curves of the PCR products obtained with the SNP primer pairs, it becomes evident that some cultivars share the same DNA sequence even though their melting curves differ. An example would be cultivar Elmo with primer pair SNP1. Even though the melting curve produced by Elmo's amplicons deviates from the cluster located to its left it shares the same DNA sequence (Figures 63, 69). The other obtained results for primer pair SNP1 match with the melting curves. Olga and Katalin share the same sequence and the same applies to Jafar and Maragiá. Analisto has a unique sequence and the melting curve produced by its PCR product is also separated from the rest. Olga and Katalin, with only one G/G SNP, melt at a lower temperature than Elmo and its cluster, which contain two G/G SNPs.

Sample	Cultivar	Type	Date	Sequence
CP51	Cicerone	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP52	Vulcano	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP53	Flamenco	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP55	Elmo	Desi	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP56	Rondo	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP57	Donia	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP58	Dora	Gulabi	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP59	Maragiá	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP59	Maragiá	Kabuli	15.05.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP510	Analisto	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP510	Analisto	Kabuli	15.05.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP511	Jafar	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP511	Jafar	Kabuli	15.05.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP512	Gavdos	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP512	Gavdos	Kabuli	15.05.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP513	Boni	Desi	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP514	Sultano	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP515	Reale	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP515	Reale	Kabuli	15.05.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP516	Amorgos	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP517	Thiva	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP518	CDC Orion	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP519	Elvir	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP519	Elvir	Kabuli	15.05.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP520	Sokol	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP521	Olga	Gulabi	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP521	Olga	Gulabi	15.05.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP522	Badli	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP523	Pasciá	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP524	Katalin	Desi	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP524	Katalin	Desi	15.05.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT
CP525	Twist	Kabuli	25.03.2024	ACGATTGCTAAGGAAGTCGACGTGTATCCGTGCTAAGATTAGATGGACCAACAGCAGTGAATAGGCTGCTAAATGTT

Figure 69: Sequencing results of samples amplified with primer pair SNP1. SNPs are color-coded in red (not reliable), yellow (semi-reliable) and blue (reliable).

Even though the melting curves of the products of Cicerone, Flamenco and Thiva slightly differ from the cluster to the left, they all share the same sequence when amplified with primer pair SNP2 (Figures 62, 70). Pasciá and Reale are unique in their sequence, and the melting curves produced by their amplicons are isolated from the rest. The melting curves of the products of Maragiá, Jafar and Gavdos overlap but their sequences are not matching. However, their obtained results are all red and they are not very reliable. It is assumed that they share the same sequence with CDC Orion. Since the cluster only contains SNPs with A or T they melt first followed by heterozygous Pasciá.

5 Results & discussion

The remaining cultivars melt at a higher temperature corresponding with their higher GC content.

Sample	Cultivar	Type	Date	Sequence
CP51	Cicerone	Kabuli	22.03.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP52	Vulcano	Kabuli	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP53	Flamenco	Kabuli	19.03.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP54	Flamenco	Kabuli	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP55	Elmo	Desi	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP56	Rondo	Kabuli	22.03.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP57	Donia	Kabuli	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP58	Dora	Gulabi	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP59	Maragjá	Kabuli	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP59	Maragjá	Kabuli	15.05.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP510	Anallisto	Kabuli	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP511	Jafar	Kabuli	22.03.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP511	Jafar	Kabuli	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP511	Jafar	Kabuli	15.05.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP512	Gavdos	Kabuli	22.03.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP512	Gavdos	Kabuli	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP512	Gavdos	Kabuli	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP513	Bori	Desi	22.03.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP514	Sultano	Kabuli	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP515	Reale	Kabuli	22.03.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP516	Amorgos	Kabuli	22.03.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP517	Thiva	Kabuli	22.03.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP517	Thiva	Kabuli	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP517	Thiva	Kabuli	15.05.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP518	CDC Orion	Kabuli	22.03.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP519	Elxir	Kabuli	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP520	Sokol	Kabuli	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP521	Olga	Gulabi	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP522	Badli	Kabuli	22.03.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP523	Pasciá	Kabuli	19.03.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP524	Katalin	Desi	11.04.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC
CP525	Twist	Kabuli	22.03.2024	ACTTTGCATCCCAATTTTCCTAT/TCA/ATCTCGTTCT/CCCTTGAACAATACAGCTGCCCTCTGAAC

Figure 70: Sequencing results of samples amplified with primer pair SNP2. SNPs are color-coded in red (not reliable), yellow (semi-reliable) and blue (reliable).

Amplification with primer pair SNP3 resulted in the formation of a melting curve cluster (Figure 66). Only the PCR products of Sokol and Olga show distinct melting curves, which can be separated from the rest. This is also reflected in their sequencing result as their sequence is unique (Figure 71). The PCR products of some random samples (Elmo, Dora, Gavdos and Twist) were sequenced from the melting curve cluster and they all share the same alleles. Due to the two G/G SNPs present in the cluster it melts at a higher temperature compared to Sokol and Olga.

Sample	Cultivar	Type	Date	Sequence
CP55	Elmo	Desi	25.03.2024	AGCAACA/AACAATTGCCCTTAATCTTAACACT/GCG/GACCGTTATTAATTGAGTATTTGACCTAATGAACCTCATCATCA
CP58	Dora	Gulabi	25.03.2024	AGCAACA/AACAATTGCCCTTAATCTTAACACT/GCG/GACCGTTATTAATTGAGTATTTGACCTAATGAACCTCATCATCA
CP512	Gavdos	Kabuli	25.03.2024	AGCAACA/AACAATTGCCCTTAATCTTAACACT/GCG/GACCGTTATTAATTGAGTATTTGACCTAATGAACCTCATCATCA
CP520	Sokol	Kabuli	25.03.2024	AGCAACA/CACAATTGCCCTTAATCTTAACACT/GCG/GACCGTTATTAATTGAGTATTTGACCTAATGAACCTCATCATCA
CP520	Sokol	Kabuli	15.05.2024	AGCAACA/CACAATTGCCCTTAATCTTAACACT/GCG/GACCGTTATTAATTGAGTATTTGACCTAATGAACCTCATCATCA
CP521	Olga	Gulabi	25.03.2024	AGCAACA/CACAATTGCCCTTAATCTTAACACT/GCG/GACCGTTATTAATTGAGTATTTGACCTAATGAACCTCATCATCA
CP521	Olga	Gulabi	15.05.2024	AGCAACA/CACAATTGCCCTTAATCTTAACACT/GCG/GACCGTTATTAATTGAGTATTTGACCTAATGAACCTCATCATCA
CP525	Twist	Kabuli	25.03.2024	AGCAACA/AACAATTGCCCTTAATCTTAACACT/GCG/GACCGTTATTAATTGAGTATTTGACCTAATGAACCTCATCATCA

Figure 71: Sequencing results of samples amplified with primer pair SNP3. SNPs are color-coded in red (not reliable), yellow (semi-reliable) and blue (reliable).

5.2 Single nucleotide polymorphism

The PCR products of Thiva and Dora have a slightly different melting behaviour than their cluster to the right but share the same sequence when amplified with primer pair SNP4 (Figures 65, 72). The products of Olga, Elmo and CDC Orion also share the same sequence and melting curves and, due to their higher GC content, melt at a higher temperature.

Sample	Cultivar	Type	Date	Sequence
CPS1	Cicerone	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS2	Vulcano	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS3	Flamenco	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS5	Elmo	Desi	11.04.2024	CAGTGTGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS5	Elmo	Desi	15.05.2024	CAGTGTGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS6	Rondo	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS7	Donia	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS8	Dora	Gulabi	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS8	Dora	Gulabi	15.05.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS9	Maraja	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS10	Analisto	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS11	Jaffer	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS12	Gavdos	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS13	Bori	Desi	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS14	Sultano	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS15	Reale	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS16	Amorgos	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS17	Thiva	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS17	Thiva	Kabuli	15.05.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS18	CDC Orion	Kabuli	11.04.2024	CAGTGTGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS18	CDC Orion	Kabuli	15.05.2024	CAGTGTGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS19	Elixir	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS20	Sokol	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS21	Olga	Gulabi	11.04.2024	CAGTGTGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS22	Badil	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS23	Pascia	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS24	Katalin	Desi	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT
CPS25	Twist	Kabuli	11.04.2024	CAATGTTGCCAGGAAATATCAAAATAAAATATAAAGTGGGGGATAAGCAGAGGGGTAACAGGACCTGTCTTCACTACTTGAATCCTCGT

Figure 72: Sequencing results of samples amplified with primer pair SNP4. SNPs are color-coded in red (not reliable), yellow (semi-reliable) and blue (reliable).

Primer pair SNP8 was one of the most exciting primer pairs due to its ability to differentiate kabuli types from desi and gulabi (Figure 64). The only exception is cultivar Flamenco whose melting behaviour and sequence are identical to desi and gulabi types (Figure 73). Due to their higher GC content the PCR product of Flamenco melts at a higher temperature together with the rest of the desi and gulabi samples. From the kabuli cluster, only a few randomly picked samples (Cicerone, Gavdos, Elixir and Twist) were sequenced.

Sample	Cultivar	Type	Date	Sequence
CPS1	Cicerone	Kabuli	15.05.2024	TTATACGCTTGTCAAGCTGCGTGACAAATGAGCTAATCCAGCGCAATCGGTAGAAAG
CPS3	Flamenco	Kabuli	22.03.2024	TTATACGCTTGTCAAGCTGCGTGACAAATGAGCTAATCCAGCGCAATCGGTAGAAAG
CPS5	Elmo	Desi	15.05.2024	TTATACGCTTGTCAAGCTGCGTGACAAATGAGCTAATCCAGCGCAATCGGTAGAAAG
CPS8	Dora	Gulabi	22.03.2024	TTATACGCTTGTCAAGCTGCGTGACAAATGAGCTAATCCAGCGCAATCGGTAGAAAG
CPS12	Gavdos	Kabuli	22.03.2024	TTATACGCTTGTCAAGCTGCGTGACAAATGAGCTAATCCAGCGCAATCGGTAGAAAG
CPS13	Bori	Desi	15.05.2024	TTATACGCTTGTCAAGCTGCGTGACAAATGAGCTAATCCAGCGCAATCGGTAGAAAG
CPS19	Elixir	Kabuli	22.03.2024	TTATACGCTTGTCAAGCTGCGTGACAAATGAGCTAATCCAGCGCAATCGGTAGAAAG
CPS21	Olga	Gulabi	22.03.2024	TTATACGCTTGTCAAGCTGCGTGACAAATGAGCTAATCCAGCGCAATCGGTAGAAAG
CPS24	Katalin	Desi	15.05.2024	TTATACGCTTGTCAAGCTGCGTGACAAATGAGCTAATCCAGCGCAATCGGTAGAAAG
CPS25	Twist	Kabuli	15.05.2024	TTATACGCTTGTCAAGCTGCGTGACAAATGAGCTAATCCAGCGCAATCGGTAGAAAG

Figure 73: Sequencing results of samples amplified with primer pair SNP8. SNPs are color-coded in red (not reliable), yellow (semi-reliable) and blue (reliable).

5 Results & discussion

Upon amplification with SNP7, the cultivars were divided into three clusters, which are distinct even though the melting curves of the products of some samples were scattering (Figure 67). CDC Orion forms one cluster by itself as it is the only heterozygous sample (Figure 74). The cluster to the left is formed by homozygous SNPs only containing A or T and, therefore, melts at a lower temperature. The cluster to the right all share homozygous SNPs with G and C and melt at a higher temperature.

Sample	Cultivar	Type	Date	Sequence
OPS1	Cicerone	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS2	Vulcano	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS2	Vulcano	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS3	Flamenco	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS4	Flamenco	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS5	Elmo	Desi	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS6	Rondo	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS7	Donia	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS8	Dora	Gulabi	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS8	Dora	Gulabi	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS9	Maragiá	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS9	Maragiá	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS10	Analisto	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS11	Jafar	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS11	Jafar	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS12	Gavdos	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS12	Gavdos	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS13	Bori	Desi	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS13	Bori	Desi	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS14	Sultano	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS14	Sultano	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS15	Reale	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS16	Amorgos	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS17	Thiva	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS17	Thiva	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS18	CDC Orion	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS18	CDC Orion	Kabuli	15.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS18	CDC Orion	Kabuli	15.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS18	CDC Orion	Kabuli	15.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS19	Elixir	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS19	Elixir	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS20	Sokol	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS21	Olga	Gulabi	15.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS21	Olga	Gulabi	15.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS22	Badil	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS22	Badil	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS23	Pasciá	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS24	Katalin	Desi	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT
OPS25	Twist	Kabuli	14.05.2024	CCAGTGA CAGGTTGATACAAAGAGGCAGCTGTTCCATCA/ACTTTGTGCTGGTGTGACGTAGCAACAGAACTGTCAA CAAT

Figure 74: Sequencing results of samples amplified with primer pair SNP7. SNPs are color-coded in red (not reliable), yellow (semi-reliable) and blue (reliable).

It was observed that chickpeas tend to be either homozygous or heterozygous for all tested SNPs as seen in Figure 68. A mixture between homozygous and heterozygous SNPs in the same amplicon occurred rarely among the tested sequences. However, if a mixture between the two was detected, the result was marked as red and therefore, is considered to be not very reliable.

5.3 Combination of results based on microsatellites & SNPs

The tested primer pairs targeting microsatellites could not differentiate twelve cultivar pairs whereas the primer pairs targeting SNPs could not differentiate 19 cultivar pairs. However, when combining the final results of microsatellites and SNPs only four cultivar pairs remain, that cannot be differentiated with either motif:

- Rondo & Donia
- Maragiá & Jafar
- Rondo & Amorgos
- Donia & Amorgos

Two possible reasons for this might be that the cultivars are too closely related to each other and thus show little to no differences in their DNA or a locus suitable for differentiation is still to be detected.

6 Conclusion

Multiple DNA-based assays involving PCR followed by HRM analysis were developed to differentiate 24 different chickpea cultivars. The assays either target microsatellites or multiple SNPs in the chickpea genome. DNA could be successfully extracted from all cultivars using a commercial kit. For microsatellites already established primer pairs from different papers were reevaluated and tested for their suitability for cultivar differentiation. Primer pairs targeting SNPs were newly designed and their amplified products were sequenced for a deeper understanding of the melt curves. All primer pairs listed in this thesis were preselected by their *in silico* test results and tested in the lab. While most of the tested primer pairs were suitable for cultivar differentiation, singular tested primer pairs did not contribute to the bigger picture.

Cultivar differentiation properties of singular primer pairs were depicted in excel data sheets using a color code (red - not differentiable, yellow - maybe differentiable, green - differentiable), which were then used to generate one final excel data sheet for microsatellites and SNPs including most suitable primer pairs for differentiation. All primer pairs tested for cultivar differentiation provided highly repeatable results even when using a second DNA extract. An exception are primer pairs CP14 and SNP7. While primer pair CP14 was slightly adapted, repeatability could not be increased even by adjustment of magnesium chloride concentration. Primer pair SNP7 was not adjusted in any way and despite lacking repeatability, the melting curves were repeatable enough to produce three distinct groups, which were useful for cultivar differentiation.

Amplicons containing three or more SNPs were sequenced to gain more information about the position and shape of their melting curves. Some sequencing results remained very unreliable despite multiple repetitions and are thus to be treated with caution. Overall it was observed that despite showing different melting behaviour, some chickpea cultivars share the same DNA sequence. Since the melting behaviour of the amplicons is entirely based on their DNA sequence, this assessment is rather peculiar. Additionally, it was observed that chickpea cultivars are either fully homozygous or heterozygous for all SNPs contained in a single amplicon. A mixture between homozygous and heterozygous SNPs was rarely observed, and if it occurred, the results were marked as not reliable by the sequencing device.

Most of the analyzed cultivars could be differentiated with the established assays. The tested primer pairs targeting microsatellites were unable to differentiate twelve cultivar pairs, while those targeting SNPs failed to differentiate 19 cultivar pairs. However, when combining the final results from both microsatellites and SNPs, only four cultivar pairs remained indistinguishable. Either the cultivars are too closely related to each other or a suitable locus for differentiation in their genome has yet to be detected.

Further work is still required in this research field as it is still unclear why different

melting behaviour between cultivars is observed during HRM analysis even though their DNA sequence is identical. Furthermore, unreliable sequencing results should be repeated. Additionally, further tests are necessary to investigate whether the four remaining cultivar pairs, which are yet to be differentiated, can be distinguished with additional primer pairs. Also, further steps for repeatability improvement of primer pair CP14 and SNP7 can be considered.

On top of that, further tests regarding the functionality of these assays with processed chickpeas might be interesting as, in this thesis, only raw material which was ground prior was used.

Overall, DNA-based assays with PCR followed by HRM show big potential for cultivar differentiation and provide a cheap and fast detection method for food fraud.

7 Materials

7.0.1 Laboratory equipment

Name	Manufacturer
Analytical balance Sartorius Talent TE214S	Sartorius
Centrifuge 5424	Eppendorf
Centrifuge Mini Star	VWR International
Galaxy Mini-Centrifuge	VWR International
Ikamag [®] RCT	IKA
PCR Workstation Pro	VWR International
Pipette Xplorer, 0.5–10 µL	Eppendorf
Pipette Xplorer, 1–20 µL	Eppendorf
Pipette Xplorer, 10–200 µL	Eppendorf
Pipette Xplorer, 50–1000 µL	Eppendorf
PowerPac HV Power Supply	Bio-Rad Laboratories
PyroMark [®] Q48	Qiagen
QuantStudio [™] 5 Thermocycler	Thermo Fisher
Qubit [™] 4 Fluorometer	Thermo Fisher
Rotor-Gene Q Thermocycler	Qiagen
Sub-Cell GT Gel Caster	Bio-Rad Laboratories
Sub-Cell [®] GT MP comb	Bio-Rad Laboratories
Thermomixer C Eppendorf	Eppendorf
Transilluminator UVT-20M	Herolab
Vortexer Vortex-Genie 2	Scientific Industries
Vortexer Classic Advanced Vortex Mixer	VELP Scientifica

7.0.2 Software

Name	Developer
LaTeX	LaTeX project team
Office 365	Microsoft
OriginPro 2024 10.1.0.170	OriginLab
PyroMark [®] Assay Design 2.0.1.15	Qiagen
PyroMark [®] Q48 Autoprep 2.4.2	Qiagen
QuantStudio [™] Design & Analysis Software 1.4.3	Thermo Fisher
Rotor-Gene Q Series Software 2.3.1	Qiagen

7.0.3 Solutions & kits

Name	Manufacturer
Agarose High Resolution	Carl Roth
CoralLoad Concentrate 10x	Qiagen
DNA-ExitusPlus™ IF	PanReac AppliChem ITW Reagents
ELGA water	Department of Analytical Chemistry University of Vienna (WG Köllensperger)
GelRed® Nucleic Acid Gel Stain	Biotium
Low Molecular Weight DNA Ladder	New England BioLabs GmbH
NucleoSpin Plant II Kit	Macherey-Nagel
Primer	Sigma-Aldrich
Qubit™ 1x ds DNA BR Assay Kit	Invitrogen Thermo Fisher Scientific
RNase A (Ribonuclease A)	Macherey-Nagel
RNase free water	Qiagen
ROTI®Load DNA-orange 1 (with Glycerol)	Carl Roth GmbH + Co. KG
10x TBE Buffer	Invitrogen
2x Type-It® HRM PCR Kit	Qiagen

7.0.4 Webserver

Name	Link
Chemical Sketch Tool	https://www.rcsb.org/chemical-sketch
NCBI - Nucleotide BLAST	https://blast.ncbi.nlm.nih.gov/Blast.cgi
OligoAnalyzer	https://www.idtdna.com/pages/tools/oligoanalyzer
OligoCalc	http://biotools.nubic.northwestern.edu/OligoCalc.html
RNAfold Webserver	http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi

7.0.5 Other utensils

Name	Manufacturer
duoSHIELD™ PFT Nitrile 240 gloves	SHIELD Scientific
Eppendorf Tubes® (200 µL, 1.5 mL, 2 mL, 5 mL)	Eppendorf
Filter tips (10 - 1000 µL)	Sarstedt
PCR Tubes (0.2 mL)	Thermo Fisher
PCR TW 8-Tube Strip (0.2 mL)	Biozym
Strip Tubes and Caps (0.1 mL)	Qiagen
Qubit™ Assay Tubes	Invitrogen

Bibliography

- [1] M. Kumar and K. Rani, 'Epigenomics in stress tolerance of plants under the climate change,' *Molecular Biology Reports*, vol. 50, no. 7, pp. 6201–6216, Jun. 2023. DOI: [10.1007/s11033-023-08539-6](https://doi.org/10.1007/s11033-023-08539-6).
- [2] K. Calvin, D. Dasgupta, G. Krinner *et al.*, *IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland*. P. Arias, M. Bustamante, I. Elgizouli *et al.*, Eds. Jul. 2023. DOI: [10.59327/ipcc/ar6-9789291691647](https://doi.org/10.59327/ipcc/ar6-9789291691647).
- [3] D. Jain and D. Chattopadhyay, 'Analysis of gene expression in response to water deficit of chickpea (*cicer arietinum* l.) varieties differing in drought tolerance,' *BMC Plant Biology*, vol. 10, no. 1, p. 24, 2010. DOI: [10.1186/1471-2229-10-24](https://doi.org/10.1186/1471-2229-10-24).
- [4] B. S. Chandana, R. K. Mahto, R. K. Singh *et al.*, 'Epigenomics as potential tools for enhancing magnitude of breeding approaches for developing climate resilient chickpea,' *Frontiers in Genetics*, vol. 13, Jul. 2022. DOI: [10.3389/fgene.2022.900253](https://doi.org/10.3389/fgene.2022.900253).
- [5] H. F. Erbersdobler, C. A. Barth and G. Jahreis, 'Legumes in human nutrition. nutrient content and protein quality of pulses.,' *Ernahrungs Umschau*, no. 64(9), pp. 134–139, Sep. 2017. DOI: [10.4455/eu.2017.034](https://doi.org/10.4455/eu.2017.034).
- [6] van der Maesen LJG, 'Origin, history and taxonomy of chickpea,' *Saxena MC, Singh KB (eds) The chickpea*, pp. 11–34, 1987.
- [7] H. Vural and A. Akcin, 'Molecular analysis of chickpea species through molecular markers,' *Biotechnology & Biotechnological Equipment*, vol. 24, no. 2, pp. 1828–1832, Jan. 2010. DOI: [10.2478/v10133-010-0046-y](https://doi.org/10.2478/v10133-010-0046-y).
- [8] S. Amini, R. Maali-Amiri, S.-S. Kazemi-Shahandashti *et al.*, 'Effect of cold stress on polyamine metabolism and antioxidant responses in chickpea,' *Journal of Plant Physiology*, vol. 258–259, p. 153387, Mar. 2021. DOI: [10.1016/j.jplph.2021.153387](https://doi.org/10.1016/j.jplph.2021.153387).
- [9] A. K. Jukanti, P. M. Gaur, C. L. L. Gowda and R. N. Chibbar, 'Nutritional quality and health benefits of chickpea (*cicer arietinum* l.): A review,' *British Journal of Nutrition*, vol. 108, no. S1, S11–S26, Aug. 2012. DOI: [10.1017/s0007114512000797](https://doi.org/10.1017/s0007114512000797).
- [10] U. Khalid, M. Q. Waheed, N. Parveen, M. A. R. Arif and A. Arif, 'Estimation of genetic diversity using seed storage protein (ssp) profiling in wild and cultivated species of *Cicer* l,' *Molecular Biology Reports*, vol. 50, no. 5, pp. 4175–4185, Mar. 2023. DOI: [10.1007/s11033-023-08358-9](https://doi.org/10.1007/s11033-023-08358-9).

- [11] S. S. Li, C. W. Kendall, R. J. de Souza *et al.*, 'Dietary pulses, satiety and food intake: A systematic review and meta-analysis of acute feeding trials,' *Obesity*, vol. 22, no. 8, pp. 1773–1780, May 2014. DOI: [10.1002/oby.20782](https://doi.org/10.1002/oby.20782).
- [12] D. Zohary, M. Hopf and E. Weiss, *Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin*. Oxford University Press, Mar. 2012, ISBN: 9780199549061. DOI: [10.1093/acprof:osobl/9780199549061.001.0001](https://doi.org/10.1093/acprof:osobl/9780199549061.001.0001).
- [13] G. Ladizinsky and A. Adler, 'The origin of chickpea *cicer arietinum* l.,' *Euphytica*, vol. 25, no. 1, pp. 211–217, Jan. 1976. DOI: [10.1007/bf00041547](https://doi.org/10.1007/bf00041547).
- [14] U. Basu, R. Srivastava, D. Bajaj *et al.*, 'Genome-wide generation and genotyping of informative snps to scan molecular signatures for seed yield in chickpea,' *Scientific Reports*, vol. 8, no. 1, Sep. 2018. DOI: [10.1038/s41598-018-29926-1](https://doi.org/10.1038/s41598-018-29926-1).
- [15] P. Ghaffari, R. Talebi and F. Keshavarzi, 'Genetic diversity and geographical differentiation of iranian landrace, cultivars, and exotic chickpea lines as revealed by morphological and microsatellite markers,' *Physiology and Molecular Biology of Plants*, vol. 20, no. 2, pp. 225–233, Mar. 2014. DOI: [10.1007/s12298-014-0223-9](https://doi.org/10.1007/s12298-014-0223-9).
- [16] T. T. Nguyen, P. W. J. Taylor, R. J. Redden and R. Ford, 'Genetic diversity estimates in *Cicer* using aflp analysis,' *Plant Breeding*, vol. 123, no. 2, pp. 173–179, Apr. 2004. DOI: [10.1046/j.1439-0523.2003.00942.x](https://doi.org/10.1046/j.1439-0523.2003.00942.x).
- [17] FAO, *Faostat* — fao.org, <https://www.fao.org/faostat/en/#data/QCL>, [Accessed 02-07-2024].
- [18] P. Joshi, 'Genetic variability, heritability and genetic advance study for seed yield and yield component traits in a chickpea recombinant inbred line (ril) population,' *International Journal of Pure & Applied Bioscience*, vol. 6, no. 2, pp. 136–141, Mar. 2018. DOI: [10.18782/2320-7051.6231](https://doi.org/10.18782/2320-7051.6231).
- [19] *Chickpea - the global bean project*, <https://www.globalbean.eu/publications/chickpea/>, [Accessed 04-07-2024].
- [20] P. Madesis, I. Ganopoulos, A. Anagnostis and A. Tsaftaris, 'The application of bar-hrm (barcode dna-high resolution melting) analysis for authenticity testing and quantitative detection of bean crops (leguminosae) without prior dna purification,' *Food Control*, vol. 25, no. 2, pp. 576–582, Jun. 2012. DOI: [10.1016/j.foodcont.2011.11.034](https://doi.org/10.1016/j.foodcont.2011.11.034).
- [21] T. Eker, D. Sari, H. Sari, H. S. Tosun and C. Toker, 'A kabuli chickpea ideotype,' *Scientific Reports*, vol. 12, no. 1, Jan. 2022. DOI: [10.1038/s41598-022-05559-3](https://doi.org/10.1038/s41598-022-05559-3).
- [22] M.-T. Moreno and J. I. Cubero, 'Variation in *cicer arietinum* l.,' *Euphytica*, vol. 27, no. 2, pp. 465–485, Jun. 1978. DOI: [10.1007/bf00043173](https://doi.org/10.1007/bf00043173).
- [23] K. K. Agrawal, U. P. S. Bhadauria and S. Jain, 'Effect of thermal regimes on crop growth, development and seed yield of chickpea (*cicer arietinum* l.),' in *Challenges and Opportunities in Agrometeorology*. Springer Berlin Heidelberg, 2011, pp. 499–503, ISBN: 9783642193606. DOI: [10.1007/978-3-642-19360-6_37](https://doi.org/10.1007/978-3-642-19360-6_37).

Bibliography

- [24] *Food fraud / knowledge for policy*, https://knowledge4policy.ec.europa.eu/food-fraud-quality/topic/food-fraud_en, [Accessed 04-07-2024].
- [25] *Regulation - 178/2002 - en - eur-lex — data.europa.eu*, <http://data.europa.eu/eli/reg/2002/178/oj>, [Accessed 04-07-2024].
- [26] FDA, *Economically motivated adulteration (food fraud) — fda.gov*, <https://www.fda.gov/food/compliance-enforcement-food/economically-motivated-adulteration-food-fraud>, [Accessed 02-07-2024].
- [27] R. A. Graybosch, ‘The grain crops: An overview,’ in *Encyclopedia of Food Grains*, Elsevier, 2016, p. 51.
- [28] N. K. Sethy, B. Shokeen, K. J. Edwards and S. Bhatia, ‘Development of microsatellite markers and analysis of intraspecific genetic variability in chickpea (*cicer arietinum* l.),’ *Theoretical and Applied Genetics*, vol. 112, no. 8, pp. 1416–1428, Mar. 2006. DOI: [10.1007/s00122-006-0243-0](https://doi.org/10.1007/s00122-006-0243-0).
- [29] S. Parween, K. Nawaz, R. Roy *et al.*, ‘An advanced draft genome assembly of a desi type chickpea (*cicer arietinum* l.),’ *Scientific Reports*, vol. 5, no. 1, Aug. 2015. DOI: [10.1038/srep12806](https://doi.org/10.1038/srep12806).
- [30] P. Winter, T. Pfaff, S. M. Udupa *et al.*, ‘Characterization and mapping of sequence-tagged microsatellite sites in the chickpea (*cicer arietinum* l.) genome,’ *Molecular and General Genetics MGG*, vol. 262, no. 1, pp. 90–101, Aug. 1999. DOI: [10.1007/s004380051063](https://doi.org/10.1007/s004380051063).
- [31] D. Tautz and M. Renz, ‘Simple sequences are ubiquitous repetitive components of eukaryotic genomes,’ *Nucleic Acids Research*, vol. 12, no. 10, pp. 4127–4138, 1984. DOI: [10.1093/nar/12.10.4127](https://doi.org/10.1093/nar/12.10.4127).
- [32] W. Choumane, P. Winter, F. Weigand and G. Kahl, ‘Conservation and variability of sequence-tagged microsatellite sites (stmss) from chickpea (*Cicer arietinum* l.) within the genus *Cicer*,’ *Theoretical and Applied Genetics*, vol. 101, no. 1–2, pp. 269–278, Jul. 2000. DOI: [10.1007/s001220051479](https://doi.org/10.1007/s001220051479).
- [33] A. J. Brookes, ‘The essence of snps,’ *Gene*, vol. 234, no. 2, pp. 177–186, Jul. 1999. DOI: [10.1016/s0378-1119\(99\)00219-x](https://doi.org/10.1016/s0378-1119(99)00219-x).
- [34] R. Sachidanandam, D. Weissman, S. C. Schmidt *et al.*, ‘A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms,’ *Nature*, vol. 409, no. 6822, pp. 928–933, Feb. 2001. DOI: [10.1038/35057149](https://doi.org/10.1038/35057149).
- [35] A. Vignal, D. Milan, M. SanCristobal and A. Eggen, ‘A review on snp and other types of molecular markers and their use in animal genetics,’ *Genetics Selection Evolution*, vol. 34, no. 3, May 2002. DOI: [10.1186/1297-9686-34-3-275](https://doi.org/10.1186/1297-9686-34-3-275).
- [36] M. Erali, K. V. Voelkerding and C. T. Wittwer, ‘High resolution melting applications for clinical laboratory medicine,’ *Experimental and Molecular Pathology*, vol. 85, no. 1, pp. 50–58, Aug. 2008. DOI: [10.1016/j.yexmp.2008.03.012](https://doi.org/10.1016/j.yexmp.2008.03.012).

- [37] A. W. Khan, M. Thudi, R. K. Varshney and D. Edwards, 'Sequencing the chickpea genome,' in *The Chickpea Genome*. Springer International Publishing, 2017, pp. 117–123, ISBN: 9783319661179. DOI: [10.1007/978-3-319-66117-9_10](https://doi.org/10.1007/978-3-319-66117-9_10).
- [38] B. Alberts, *Molecular biology of the cell*, 6th ed. Boca Raton, FL: CRC Press, Dec. 2014, pp. 452, 473.
- [39] D. Beyene, 'The state of knowledge in designing primer of interest for successful polymerase chain reaction: A review,' *Advanced Journal of Biological Sciences Research*, vol. 2, no. 3, pp. 1–12, Nov. 2014.
- [40] M. Słomka, M. Sobalska-Kwapis, M. Wachulec, G. Bartosz and D. Strapagiel, 'High resolution melting (hrm) for high-throughput genotyping—limitations and caveats in practical case studies,' *International Journal of Molecular Sciences*, vol. 18, no. 11, p. 2316, Nov. 2017. DOI: [10.3390/ijms18112316](https://doi.org/10.3390/ijms18112316).
- [41] R. Matissek, M. Fischer and G. Steiner, *Lebensmittelanalytik* (Springer-Lehrbuch), 6th ed. Wiesbaden, Germany: Springer, Aug. 2018, pp. 8, 93, 178–183, 199–206, 209–212, 218–220.
- [42] Macherey-Nagel, *Genomic dna from plant user manual nucleospin® plant ii midi / maxi*, Macherey-Nagel GmbH & Co. KG, Mountain View, California, USA, 2014.
- [43] P. K. S. Chan, 'Evaluation of extraction methods from paraffin wax embedded tissues for pcr amplification of human and viral dna,' *Journal of Clinical Pathology*, vol. 54, no. 5, pp. 401–403, May 2001. DOI: [10.1136/jcp.54.5.401](https://doi.org/10.1136/jcp.54.5.401).
- [44] B. Bruijns, T. Hoekema, L. Oomens, R. Tiggelaar and H. Gardeniers, 'Performance of spectrophotometric and fluorometric dna quantification methods,' *Analytica*, vol. 3, no. 3, pp. 371–384, Sep. 2022. DOI: [10.3390/analytica3030025](https://doi.org/10.3390/analytica3030025).
- [45] Carl Roth, *Sicherheitsdatenblatt ethidiumbromid $\geq 98\%$* , [Accessed 11-06-2024], 2024. [Online]. Available: <https://www.carlroth.com/medias/SDB-7870-CH-DE.pdf?context=bWFzdGVyfHhN1Y3VyaXR5RGF0YXNoZWV0c3wzMzg2MDF8YXBwbGljYXRpb24vcGRmfGhjMy9oOWIvOTEONzM5NzQwNjc1MC9TREJfNzg3MF9DSF9ERS5wZGZ8YjUwODM1NDA4MzY4N2JmNWFiYjIxYzNkMGlyYzU1YjVlYjAzMTliNjViMjQ3MmIxNGUyYmNlN2E5Njc3YTQwMg>.
- [46] ThermoFisher Scientific, *Pcr cycling parameters-six key considerations for success: Thermo fisher scientific - us*, [Accessed 17-06-2024], 2024. [Online]. Available: <https://www.thermofisher.com/at/en/home/life-science/cloning/cloning-learning-center/invitrogen-school-of-molecular-biology/pcr-education/pcr-reagents-enzymes/pcr-cycling-considerations.html>.
- [47] ThermoFisher Scientific, *Multiplex vs. singleplex qpcr: What do i need to know?* [Accessed 17-06-2024], 2022. [Online]. Available: <https://www.thermofisher.com/blog/behindthebench/multiplex-vs-singleplex-qpcr-what-do-i-need-to-know/>.

Bibliography

- [48] Biotium Glowing Products for Science™, *Evagreen® dye for qpcr & other applications*, [Accessed 17-06-2024], 2024. [Online]. Available: <https://biotium.com/technology/pcr-dna-amplification/evagreen-dye-for-qpcr/>.
- [49] Carl Roth, *Voluntary safety information roti®load dnastain 2 sybr®green, 6x conc., dnase-free, ready-to-use*, [Accessed 12-06-2024], 2021. [Online]. Available: <https://www.carlroth.com/medias/SDB-1CN6-IE-EN.pdf?context=bWFzdGVyfHN1Y3VyaXR5RGFOYXNoZWV0c3wyMzkyMzF8YXBwbGljYXRpb24vcGRmfHN1Y3VyaXR5RGFOYXNoZWV0cy9oMDIvaDQ0LzkwMjQ4Nzc1OTI2MDYucGRmfDY3MzM4ZDM2YzM4ZGNiNmM2ZTc2MmFiZGU3NWw4MDA0NDY2YzhhMWRmYTlkMWYxOGRmN2RjOTY5YWVkb0ZkMWI>.
- [50] Carl Roth, *Sicherheitsdatenblatt eva green® fluoreszenzfarbstoff 100x konz., für die elektrophorese*, [Accessed 12-06-2024], 2020. [Online]. Available: <https://www.carlroth.com/medias/SDB-8905-BE-DE.pdf?context=bWFzdGVyfHN1Y3VyaXR5RGFOYXNoZWV0c3wyMzQ0Mz18YXBwbGljYXRpb24vcGRmfHN1Y3VyaXR5RGFOYXNoZWV0cy9oYTUvaDFkLzg5OTIzMjc2MDM4NzAucGRmfGU4ODcyMWI1YzExMzkyZGU1NDAYODk4ZjZiMTA4MjUxZjZjZjA0MmMxYTI2ODVjMjExZDZlNmUzZmE4NDZhODc>.
- [51] ThermoFisher Scientific, *Essentials of real-time pcr*, [Accessed 17-06-2024], 2024. [Online]. Available: <https://www.thermofisher.com/at/en/home/life-science/pcr/real-time-pcr/real-time-pcr-learning-center/real-time-pcr-basics/essentials-real-time-pcr.html>.
- [52] T. Stellberger, A. Ehrhardt, L. Gerdes, P. Guertler and A. Baiker, ‘Droplet digital pcr (ddpcr) in der virusanalytik,’ *Schriftenreihe Gentechnik für Umwelt und Verbraucherschutz*, vol. 9, Nov. 2015.
- [53] Qiagen, *Rotor-gene® type-it® hrm® discovery handbook*, Qiagen, Germantown, Maryland, USA, 2012.
- [54] R. A. Palais, M. A. Liew and C. T. Wittwer, ‘Quantitative heteroduplex analysis for single nucleotide polymorphism genotyping,’ *Analytical Biochemistry*, vol. 346, no. 1, pp. 167–175, Nov. 2005. DOI: [10.1016/j.ab.2005.08.010](https://doi.org/10.1016/j.ab.2005.08.010).
- [55] M. Ronaghi, ‘Pyrosequencing sheds light on dna sequencing,’ *Genome Research*, vol. 11, no. 1, pp. 3–11, Jan. 2001. DOI: [10.1101/gr.150601](https://doi.org/10.1101/gr.150601).
- [56] B. Gharizadeh, M. Akhras, N. Nourizad *et al.*, ‘Methodological improvements of pyrosequencing technology,’ *Journal of Biotechnology*, vol. 124, no. 3, pp. 504–511, Jul. 2006. DOI: [10.1016/j.jbiotec.2006.01.025](https://doi.org/10.1016/j.jbiotec.2006.01.025).
- [57] Qiagen, *Pyromark® q48 autoprep user manual*, Qiagen, Germantown, Maryland, USA, 2020.
- [58] *Oligo calc: Oligonucleotide properties calculator*, [Accessed 19-06-2024], 2015. [Online]. Available: <http://biotools.nubic.northwestern.edu/OligoCalc.html>.
- [59] Institute for Theoretical Chemistry University of Vienna, *Rnafold webserver*, [Accessed 19-06-2024]. [Online]. Available: <http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>.

- [60] Integrated DNA Technologies, *OligoanalyzerTM tool*, [Accessed 19-06-2024], 2024. [Online]. Available: <https://www.idtdna.com/pages/tools/oligoanalyzer>.
- [61] National Center for Biotechnology Information, *Nucleotide blast*, [Accessed 19-06-2024]. [Online]. Available: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome.
- [62] P. Stothard, *Reverse complement*, [Accessed 19-06-2024], 2000. [Online]. Available: https://www.bioinformatics.org/sms/rev_comp.html.

Appendix

Abbreviations

- A - adenine
- AGES - Austrian Agency for Health and Food Safety
- APS - adenosine 5' phosphosulfate
- ADP - adenosine diphosphate
- AMP - adenosine monophosphate
- ATP - adenosine triphosphate
- bp - basepair
- BLAST - basic local alignment search tool
- C - cytosine
- CP - chickpea primer
- CPS - chickpea sample
- CT-value - cycle threshold value
- dATP - deoxyadenosine triphosphat
- dCTP - deoxycytidine triphosphate
- dGTP - deoxyguanosine triphosphate
- dNDP - deoxynucleotide diphosphate
- dNMP - deoxynucleotide monophosphate
- dNTPs - deoxynucleotide triphosphate
- dTTP - deoxythymidine triphosphate
- fw - forward
- G - guanine

- GC content - guanine and cytosine content
- HRM analysis - high resolution melting analysis
- LD - loading dye
- NCBI - National Center for Biotechnology Information
- NGS - next-generation sequencing
- nt - nucleotide
- NTC - no template control
- PCR - polymerase chain reaction
- PPi - pyrophosphate
- PSQ - pyrosequencing
- rev - reverse
- RFU - relative fluorescence unit
- SNP - single nucleotide polymorphism
- SSR - simple sequence repeats
- STR - short tandem repeats
- T - thymidine
- T_A - annealing temperature
- T_E - elongation temperature
- T_M - melting temperature
- TBE - Tris-borate-EDTA

List of Tables

1	Illustration of the metal plate used for QuantStudio™. The two top rows list the used annealing temperature and primer concentration for each column. CP1, CP2 and CP3 are the names of different primer pairs, which are listed in Table 6. NTC serves as the corresponding no template control.	33
2	Used recipe for gel loading buffer.	37
3	Used recipe for DNA ladder (marker).	38
4	Standard master mix recipe used for chickpeas with a primer concentration of 0.2 and 0.4 μ M. The scale is set to 52 samples (25 chickpea samples and one NTC in duplicates). Fw and rev stand for forward and reverse primer respectively.	39
5	Overview of the temperature program for PCR and HRM. Annealing temperature depends on the melting temperature of the used primer pair.	40
6	All primer pairs tested for cultivar differentiation targeting microsatellites in this thesis including their original name and source as well as their sequence, their used concentration (Conc.), annealing temperature (T_A) and whether they were suitable for differentiation (Y) or not (N). For primer pairs marked as (N) the melting curves of all cultivars were overlapping.	46
7	List of every sample loaded onto the gel depicted in Figure 40. The PCR conditions include the used primer pair, as well as the annealing temperature (T_A). Elongation temperature was uniformly set to 72 °C. A primer concentration of 0.4 μ M was used for every listed primer pair.	52
8	List of every sample loaded onto the gel depicted in Figure 44. The PCR conditions include the used primer pair, as well as the primer concentration, annealing temperature (T_A) and elongation temperature (T_E).	56
9	All primer pairs tested for cultivar differentiation targeting SNPs in this thesis including their sequence, their used concentration (Conc.), annealing temperature (T_A) and whether they were suitable for differentiation (Y) or not (N). For primer pairs marked as (N) the melting curves of all cultivars were overlapping.	67
10	First DNA extract: Sample weight and DNA concentration of all extracts. The table also includes the obtained RFU from the standard solutions used for calibration.	100
11	Second DNA extract: Sample weight and DNA concentration of all extracts. The table also includes the obtained RFU from the standard solutions used for calibration.	101

List of Figures

1	Chemical structure of CTAB drawn using Chemical Sketch Tool.	7
2	Chemical structure of SDS drawn using Chemical Sketch Tool.	8
3	Simplified schematic of the result of an agarose gel electrophoresis. Purple lines represent DNA of various sizes. First and last pocket (lane) of the gel are used for the DNA ladder. The other lanes are filled with sample DNA. Unspecific products and primer dimers are marked with arrows. . .	10
4	Simplified representation of one PCR cycle.	11
5	Simplified schematic of an amplification curve generated from a DNA sample, as well as a no template control (NTC). The threshold is marked as a red line. Graphic modified from Stellberger <i>et al.</i> [52].	13
6	This figure shows a standard melting curve generated by plotting the normalized fluorescence signal against the temperature. While the DNA is double-stranded, a high fluorescence signal is detected due to the bound fluorescent dye EvaGreen®. By continuously increasing the temperature the DNA denatures, thus leading to a decrease in fluorescence signal as EvaGreen® can no longer bind to single-stranded DNA.	15
7	This figure shows an alternative depiction of melting curves in the form of melt peaks. Melt peaks are formed by plotting the negative derivative of the fluorescence signal against the temperature. The peak maximum corresponds to the turning point of the curve.	16
8	Melting curves and melt peaks of two different chickpea cultivars amplified with primer pair SNP2. Samples were analyzed in duplicates. The sample Rondo (green) is homozygous and shows a symmetrical melting curve as well as one symmetrical melt peak. The sample Pasciá (blue) is heterozygous with an asymmetrical melting curve and a split melt peak.	17
9	Simplified depiction of the chemical reactions involved during pyrosequencing. Graphic modified from [57].	19
10	List of all cultivar names sorted by chickpea type including their source and land of origin. For experiments sample numbers CPS1 - CPS25 were used. CPS stands for chickpea sample.	22
11	Picture of the samples Analisto, Elixir and Sokol (from left to right) representing the kabuli type.	23
12	Picture of the samples Elmo, Bori and Katalin (from left to right) representing the desi type.	23
13	Picture of the samples Dora and Olga (from left to right) representing the gulabi type.	23
14	Screenshot of the website Oligo Calc. In the Nucleotide base codes text field, the primer sequence is inserted. The website calculates primer length, GC content and melting temperature. Important values are circled in red.	24

Bibliography

15	Screenshot of the website RNAfold. In the Sequence Input text field, the primer sequence is inserted. In the section Energy Parameters the DNA model is selected, and the temperature is set to 37°C and -5°C of the primer-specific melting temperature.	25
16	Screenshot of the generated graph by the website RNAfold. The color scale ranging from red to blue, where red equals 1 and blue 0, illustrates the probability of secondary structure formation. Red-colored secondary structures have a high chance of occurring, whereas blue secondary structures have a rather low chance.	25
17	Screenshot of the website OligoAnalyzer. In the Sequence text field, the primer sequence is inserted. With the buttons on the right, the probability of homo and hetero dimer formation can be calculated.	26
18	Screenshot of hetero-dimer-analysis. To analyze the formation of hetero dimers, the sequence of the second primer can be inserted.	26
19	Screenshot of the result of the homo-dimer analysis. The results are listed ranging from highest to lowest probable dimer formation. Results show the released energy by dimer formation as well as the number of matching base pairs.	27
20	Screenshot of the result of the hetero-dimer analysis. The results are listed ranging from highest to lowest probable dimer formation. Results show the released energy by dimer formation as well as the number of matching base pairs.	28
21	Screenshot of Nucleotide BLAST on the NCBI website. Primer sequence is added in the text field in the top left corner. RefSeq Genome Database and organism name (<i>Cicer arietinum</i>) are added. The program is selected for highly similar sequences.	29
22	Screenshot of the NCBI search result. Results are listed from the highest to the lowest matching sequences.	30
23	Screenshot of the highest matching result from the NCBI search. Circled in red is the number of matching base pairs. The graphics function can be used to look for the motif following the primer sequence.	30
24	Screenshot of the graphics tool for the highest matching result. The tools bar can be used to search for the primer sequence, flip DNA strands, set markers and change the sequence text view to simplify the search for the microsatellite.	30
25	Screenshot of the graphics tool for the highest matching result. The green field shows the primer sequence. A blue marker is placed at the 5' end of the primer sequence.	31
26	Screenshot of Sequence Text View feature. The microsatellite is circled in red.	31

27	Screenshot of the QuantStudio Design & Analysis Software. Primer pairs CP1 (green), CP2 (red) and CP3 (blue) tested in the concentrations 0.2 and 0.4 μ M all generate one singular peak. The no template control (NTC) of primer pair CP2 contains a primer dimer. The dimer peak is marked with an arrow.	32
28	Simplified schematic representation of DNA extraction steps using the NucleoSpin TM Plant II kit by Macherey-Nagel TM . Purple and green lines represent two different types of columns used in the kit.	35
29	Screenshot of the Rotor-Gene Q Series Software of the amplification curve of all samples with primer pair CP8.	41
30	Screenshot of the Rotor-Gene Q Series Software during the normalization step of primer pair CP8. The borders can be adjusted manually.	42
31	First evaluation draft for primer pair CP8. The first row and column list samples 1 to 25 in ascending order.	42
32	Second evaluation of primer pair CP8. The first row and column list samples 1 to 25 in ascending order.	43
33	Screenshot of the result of a sequencing run depicted in the PyroMark Q48 Autoprep software. The picked result shows a sequence containing all three listed color codes. The upper section of the screenshot shows the analyzed samples including the primer pair and the well number. The middle section lists the recorded sequence of the selected samples. Each peak represents an inserted nucleotide, and the intensity of said peak is proportional to the number of inserted nucleotides of the same type. The gray fields indicate the locations of the SNPs. The bottom section is a histogram representing the predetermined sequence, which is selected prior to the sequencing run. The gray areas represent the SNP locations and indicate possible sequencing outcomes.	44
34	Overview of the <i>in silico</i> test results of primer pair CP8 including primer sequence, length, GC content, melting temperature (T_M), repetitive motif (Rep), amplicon length, secondary structures (Sec) and potentially formed dimers (Dim). At the bottom, the location of the forward and reverse primer and the microsatellite are marked with yellow and cyan, respectively.	45
35	The left side shows the melting curves of the amplicons of all 25 chickpea samples analyzed in duplicates using primer pair CP14. Identical samples share the same color. The section between 79 and 83 $^{\circ}$ C seemed promising for cultivar differentiation, but the repeatability of the results was rather low. On the right the two melt peaks are depicted. Varying peak intensities between the different samples can be observed.	47

36	At the top, all the melting curves (left) and melt peaks (right) of the products of the handpicked samples analyzed in duplicates using primer pair CP14 are depicted. Identical samples share the same color. The middle row shows the melting curves (left) and melt peaks (right) of the products of samples 2, 3, 15, 16, and 20, which showed low repeatability in the prior experiment. While the scattering was slightly reduced, it can still be observed in this graphic. The bottom row shows melting curves (left) and melt peaks (right) of the products of samples 13, 14, 18, 21 and 23, which showed good repeatability in the prior experiment. While 13 and 18 remained the same, 24, 21 and 23 were scattered.	48
37	The left side shows the melting curves of the products of all 25 chickpea samples analyzed in duplicates using primer pair CP14a. Identical samples share the same color. While the repeatability could still be improved, the melting curve is more complex compared to CP14b and might provide more potential for cultivar differentiation. On the right the two melt peaks are depicted. Varying peak intensity between the products of the different samples and a shift for the peak maximum of the second peak between individual samples can be observed.	49
38	The left side shows the melting curves of the products of all 25 chickpea samples analyzed in duplicates using primer pair CP14b. Identical samples share the same color. While the repeatability increased slightly in comparison to primer pair CP14, the melting curves of the products of all samples look highly similar. On the right the melt peak is depicted. The peak intensity between the products of all samples differs slightly but not as much as for products obtained with CP14a. A slight shift for the peak maximum between individual samples can be observed.	49
39	Amplification curves of all 25 chickpea samples analyzed in duplicates using primer pair CP14a. Identical samples share the same color. The chickpea samples with the highest CT-value Rondo (CPS6) and Amorgos (CPS16) are marked with arrows.	50
40	Picture of an agarose gel used to separate DNA bands of different sizes. Numbers 1 to 30 indicate the different lanes. The first and last lane contain the DNA ladder with the size of the individual bands indicated on the left side. The red boxes encircle samples, which were taken from the same PCR assay. For further information regarding the content of the lanes view Table 7.	51
41	On the left, the melting curves of the products of CPS13 (first and second DNA extracts) with primer pair CPS14a from two separate PCR runs are shown. The right side shows the melting curves of the products of samples CPS3 and CPS4 as well as CPS5 (first and second DNA extracts) with primer pair CPS14a from two separate PCR runs. Identical samples share the same color. The melting curves scattered even though the experimental conditions were identical.	53

42	On the left a screenshot of the results of primer pair CP16 in the QuantStudio™ Design & Analysis software is depicted. Results are not repeatable. The right side shows the results of the conducted gel-electrophoresis. Lane one contains the DNA ladder and lanes two to six contain chickpea DNA amplified with CP16.	54
43	On the left, all melting curves of the products of all samples with primer pair CP20 are depicted. Two different DNA extracts analyzed with two different PCR runs were compared. For each sample four melting curves are shown as the first and the second extract were analyzed in duplicates. Identical samples share the same color. On the right, a few samples whose amplicons produced different melting curve shapes were selected from the graphic on the left to showcase the good repeatability between the two extracts and different PCR runs.	54
44	Picture of an agarose gel used to separate DNA bands of different sizes. Numbers 1 to 30 indicate the different lanes. The first and last lane contain the DNA ladder with the size of the individual bands indicated on the left side. The red boxes encircle samples, which were taken from the same PCR assay. For further information regarding the content of the lanes view Table 8.	55
45	The top row shows amplification curves (left) and melting curves (right) generated from the products of all chickpea samples analyzed in duplicates with primer pair CP7. A primer concentration of 0.6 μ M and an annealing temperature of 72 °C were used. The bottom row shows the amplification curves (left) and melting curves (right) with primer pair CP7 and optimized conditions. The primer concentration and annealing temperature were reduced to 0.2 μ M and 68 °C respectively. Identical samples share the same color.	57
46	Melting curves of the products of cultivars Elmo and Twist amplified with primer pair CP8. Two DNA extracts were tested in two separate PCR runs and their melting curves were combined in this figure. A slight shift to the left can be observed for the second DNA extracts.	58
47	Representation of the melting curves of the products of two sample pairs with the first and the second DNA extracts which would be marked as green (differentiable) by the newly set criterion. Their melting curves differ by 0.15 °C. Identical samples share the same color. Samples were analyzed with primer pair CP8 and optimized settings as listed in Table 6.	59
48	Representation of the melting curves of the products of three sample pairs with the first and the second DNA extracts which would be marked as red (not differentiable) by the newly set criterion. Their melting curves differ by less than 0.1 °C. Identical samples share the same color. Samples were analyzed with primer pair CP8 and optimized settings as listed in Table 6.	59

49	Melting curves of the first and the second DNA extracts of the products of chickpea cultivar Analisto with primer pair CP8. The first and the second extracts were diluted 1:100 or to 2.5 ng/ μ L. Regardless of their initial DNA concentration, the melting curves overlap.	60
50	By comparing the excel sheets of different primer pairs and grouping complementary ones a final excel sheet for microsatellites is created. Samples are sorted by their type, starting with kabuli, followed by desi and gulabi. All but twelve cultivar pairs could be differentiated using the listed primer pairs in the legend.	61
51	Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair CP8. Identical samples share the same color.	62
52	Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair CP14a. Identical samples share the same color.	62
53	Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair CP4. Identical samples share the same color.	63
54	Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair CP6. Identical samples share the same color.	63
55	Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair CP1. Identical samples share the same color.	64
56	Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair CP19. Identical samples share the same color.	64
57	Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair CP11. Identical samples share the same color.	65
58	Overview of the <i>in silico</i> test results of primer pair SNP2 including primer sequence, length, GC content, melting temperature (T_M), amplicon length, secondary structures (Sec) and potentially formed dimers (Dim). At the bottom, the location of the forward and reverse primer, and the individual SNPs are marked with yellow and cyan, respectively.	66
59	This figure shows the melting curves of the products of all 25 chickpea samples analyzed in duplicates using primer pair SNP7. Identical samples share the same color.	68
60	Melting curves of the products of all 25 chickpea samples amplified with primer pair SNP2. This figure combines the melting curves of two separate PCR runs with two different DNA extracts. The melting curves produced by the amplicons of the same sample overlap completely and no shift can be observed. Identical samples share the same color.	69
61	By comparing the excel sheets of different primer pairs and grouping complementary ones a final excel sheet for SNPs is created. Samples are sorted by their type, starting with kabuli, followed by desi and gulabi. All but 19 cultivar pairs could be differentiated using the listed primer pairs in the legend.	70
62	Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair SNP2. Identical samples share the same color.	71

63	Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair SNP1. Identical samples share the same color.	71
64	Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair SNP8. This is the only primer pair which could differentiate between different chickpea types. Identical samples share the same color.	72
65	Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair SNP4. Identical samples share the same color.	72
66	Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair SNP3. Identical samples share the same color.	73
67	Differentiation of the melting curves of the products of 24 chickpea cultivars with primer pair SNP7. Identical samples share the same color.	73
68	Melt peaks of the products of three chickpea cultivars amplified with primer pair SNP7. All three SNPs in Elixir and Dora are homozygous and show one distinct peak. Since Elixir is homozygous with A and T the melt peak is located at a lower temperature compared to Dora, whose SNPs are composed of G and C. CDC Orion is heterozygous for all three SNPs and thus two peaks of varying intensity can be observed.	74
69	Sequencing results of samples amplified with primer pair SNP1. SNPs are color-coded in red (not reliable), yellow (semi-reliable) and blue (reliable).	75
70	Sequencing results of samples amplified with primer pair SNP2. SNPs are color-coded in red (not reliable), yellow (semi-reliable) and blue (reliable).	76
71	Sequencing results of samples amplified with primer pair SNP3. SNPs are color-coded in red (not reliable), yellow (semi-reliable) and blue (reliable).	76
72	Sequencing results of samples amplified with primer pair SNP4. SNPs are color-coded in red (not reliable), yellow (semi-reliable) and blue (reliable).	77
73	Sequencing results of samples amplified with primer pair SNP8. SNPs are color-coded in red (not reliable), yellow (semi-reliable) and blue (reliable).	77
74	Sequencing results of samples amplified with primer pair SNP7. SNPs are color-coded in red (not reliable), yellow (semi-reliable) and blue (reliable).	78

Raw data

Table 10: First DNA extract: Sample weight and DNA concentration of all extracts. The table also includes the obtained RFU from the standard solutions used for calibration.

Sample	Date	Weight (g)	DNA (ng/ μ L)	RFU
Standard 1	30.10.2024		0	131.5
Standard 2	30.10.2024		100	15452.2
CPS1	30.10.2024	0.1695	82.3	
CPS2	30.10.2024	0.1630	54.3	
CPS3	30.10.2024	0.1665	41.5	
CPS4	30.10.2024	0.1521	53.8	
CPS5	30.10.2024	0.1520	53.2	
CPS6	30.10.2024	0.1548	44.4	
CPS7	30.10.2024	0.1535	59.5	
CPS8	30.10.2024	0.1504	54.1	
Standard 1	31.10.2024		0	78.3
Standard 2	31.10.2024		100	6545.1
CPS9	31.10.2024	0.1509	58.2	
CPS10	31.10.2024	0.1510	48.3	
CPS11	31.10.2024	0.1501	49.7	
CPS12	31.10.2024	0.1537	55.3	
CPS13	31.10.2024	0.1581	19.2	
CPS14	31.10.2024	0.1509	34.7	
CPS15	31.10.2024	0.1577	35.6	
CPS16	31.10.2024	0.1535	32.4	
Standard 1	13.11.2024		0	146.0
Standard 2	13.11.2024		100	15657.0
CPS17	13.11.2024	0.1510	60.2	
CPS18	13.11.2024	0.1484	37.8	
CPS19	13.11.2024	0.1553	47.1	
CPS20	13.11.2024	0.1584	41.3	
CPS21	13.11.2024	0.1588	45.6	
CPS22	13.11.2024	0.1500	40.4	
CPS23	13.11.2024	0.1528	37.0	
CPS24	13.11.2024	0.1586	21.9	
CPS25	13.11.2024	0.1487	30.6	

Table 11: Second DNA extract: Sample weight and DNA concentration of all extracts. The table also includes the obtained RFU from the standard solutions used for calibration.

Sample	Date	Student	Weight (g)	DNA (ng/ μ L)	RFU
Standard 1	04.03.2024			0	215.3
Standard 2	04.03.2024			100	17775.1
CPS1	04.03.2024	Sofie	0.1530	44.7	
CPS2	04.03.2024	Sofie	0.1580	61.5	
CPS3	04.03.2024	Sofie	0.1577	37.2	
CPS4	04.03.2024	Sofie	0.1484	37.3	
CPS5	04.03.2024	Klementina	0.1458	59.9	
CPS6	04.03.2024	Klementina	0.1565	42.1	
CPS7	04.03.2024	Klementina	0.1474	36.8	
CPS8	04.03.2024	Klementina	0.1624	35.4	
CPS9	04.03.2024	Isabel	0.1550	34.2	
CPS10	04.03.2024	Isabel	0.1459	24.1	
CPS11	04.03.2024	Isabel	0.1637	32.1	
CPS12	04.03.2024	Isabel	0.1498	34.4	
Standard 1	05.03.2024			0	138.6
Standard 2	05.03.2024			100	14544.6
CPS13	05.03.2024	Sofie	0.1560	29.5	
CPS14	05.03.2024	Sofie	0.1616	59.4	
CPS15	05.03.2024	Sofie	0.1520	49.0	
CPS16	05.03.2024	Sofie	0.1593	95.3	
CPS17	05.03.2024	Klementina	0.1467	69.1	
CPS18	05.03.2024	Klementina	0.1576	47.6	
CPS19	05.03.2024	Klementina	0.1538	60.3	
CPS20	05.03.2024	Klementina	0.1531	37.4	
CPS21	05.03.2024	Isabel	0.1581	48.5	
CPS22	05.03.2024	Isabel	0.1475	73.8	
CPS23	05.03.2024	Isabel	0.1500	78.5	
CPS24	05.03.2024	Isabel	0.1586	45.7	
CPS25	05.03.2024	Isabel	0.1437	75.3	