universität wien

DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

„Domestication and Evolution: Analyses of the Genomes of Selected Cultivated Chile Peppers"

verfasst von / submitted by

Sophie Rothen

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, 2019 / Vienna, 2019

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 190 445 344
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Lehramtsstudium UF Biologie und Umweltkunde UF Englisch
Betreut von / Supervisor:	Assoz. Prof. Mag. Dr. Hanna Schneeweiss, Privatdoz.

ERKLÄRUNG

Hiermit versichere ich,

dass ich die vorliegende Diplomarbeit selbstständig verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und mich auch sonst keiner unerlaubter Hilfe bedient habe,

dass ich dieses Diplomarbeitsthema bisher weder im In- noch im Ausland in irgendeiner Form als Prüfungsarbeit vorgelegt habe

und dass diese Arbeit mit der vom Begutachter beurteilten Arbeit vollständig übereinstimmt.

Wien,

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor, Assoz. Prof. Mag. Dr. Hanna Schneeweiss, for her continuous support and academic guidance. With endless patience and encouragement, she helped me overcome any struggles that arose during my time at the department and made working on my thesis an educational and insightful experience. Thank you for always having an open door for me and for your enthusiastic approach to research.

I am grateful to Mag. Dr. Eva Temsch for her assistance with flow cytometric measurements and for her helpful suggestions concerning data interpretation.

A very special thanks goes out to the entire Plant Evolutionary Cytogenetics research team for their encouragement and support, and for creating a warm and welcoming atmosphere at the department.

I want to thank my parents, other family members and friends for their unconditional love and support, and for being forever reassuring and encouraging about my academic endeavors.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

DSB	Double strand break
WGD	Whole genome duplication
NOR	Nucleolus organizing region
FCM	Flow Cytometry
NGS	Next generation sequencing
PI	Propidium Iodide
CV	Coefficient of variation
S.D.	Standard deviation
HCL	Haploid chromosome length

Part I – Literature Review

1. PLANT GENOME EVOLUTION

The genome consists of the genetic material of the organism, its DNA, as well as associated proteins, which are jointly organized into chromosomes and located within the nucleus of eukaryotic cells. Chromosomes store the genetic information of the organism for its correct replication and transmission, regulate gene activity and control the recombination progress (Stebbins 1971). The study of chromosomes – cytogenetics – includes analyzing chromosome number, structure, function and behavior to understand processes of genome diversification and to address evolutionary and phylogenetic questions. While conventional cytogenetical methods such as G-banding or Feulgen staining are still useful and important tools to research chromosome behavior (Stace 2000), more modern techniques in molecular cytogenetics such as fluorescent *in situ* hybridization or next-generation sequencing offer much more detailed insights into DNA sequence variations and the molecular structure of chromatin, recently focusing more on repetitive DNA sequences and their role in genome evolution (Heslop-Harrison & Schwarzacher 2011; Salmon & Ainouche 2015; Weiss-Schneeweiss et al. 2015).

In contrast to the circular chromosomes of prokaryotes, the linearity of eukaryotic chromosomes and increased genome complexity allows for much higher levels of variation in DNA amounts and is crucial for the diversification of the eukaryotic genome (Schubert 2007). Plant genomes widely differ in size, shape, and number of chromosomes, size of the genomes, as well as in the composition of DNA. All of these factors are subject to changes during evolution (Schubert 2007), and can lead to the divergence of related taxa and eventually to speciation (Rieseberg 2001; Stebbins 1971; Levin 2002). Each species/taxon has its own characteristic karyotype, its "phenotypic appearance of the somatic chromosomes in contrast to their genic contents" (Levitzky 1931, cited in Weiss-Schneeweiss & Schneeweiss 2013). Distinctive karyotypes, i.e. variations of chromosome number and structure, result from various complex chromosomal rearrangements, which include deletions, duplications, translocations, inversions, fusions and fissions, loss or gain of single chromosomes, duplications of entire genomes and changes in repetitive DNA content (Schubert 2007). Due to these rearrangements and numerical chromosomal changes, the genetic diversity of plants is striking. While the nuclear DNA content, for example, is largely consistent

within single species (Heslop-Harrison & Schwarzacher 2011), it may spectacularly vary *among* species. Genome sizes of angiosperms vary 2400-fold, from 0.063 Gb in *Genlisea G. aurea* (Greilhuber et al. 2008) to 148.8 Gb in *Paris quadrifolia* (Pellicer et al. 2010). As different mechanisms are responsible for variation in chromosome numbers and genome size, these two characters do not correlate and thus, plant species with larger genomes do not automatically have to have more chromosomes (Weiss-Schneeweiss & Schneeweiss 2013). Chromosome numbers in plants are also highly variable and range from $n = 2$ in six species of angiosperms to $n > 320$ in *Sedum suaveolens*, a succulent of the Crassulaceae family, and $n \sim 720$ in *Ophioglossum reticulatum*, a species of fern (Lysák & Schubert 2013).

In this chapter, mechanisms leading to genetic diversity, in particular structural and numerical chromosome alterations with the focus on plant evolution, are discussed. As two processes which drive genome change and evolution, polyploidy and domestication are emphasized, and the role of repetitive DNA for karyotype characterization is highlighted.

1.1. Chromosomal rearrangements

In eukaryotic genomes, gene order is conserved over wide evolutionary distances (Schmidt & Heslop-Harrison 1998). Comparative genomics often relies on synteny, i.e. “the degree to which genes remain on corresponding chromosomes” (Tang et al. 2008), as inferences can be made from simpler model organisms which are then applied to more complex genomes of related taxa. The comparison of gene arrangements provides evidence for shared ancestry of genomes as well as single chromosomes and helps understand genome evolution. While synteny and collinearity provide helpful frameworks to analyze, for example, karyotype evolution of vertebrates, their deduction is less straightforward in angiosperms due to frequent genome reshuffling, recurrent cycles of whole-genome duplications and subsequent gene loss, as well as large amounts of dispersed repetitive DNA elements. Because of the abundance in restructuring processes, even genomes of closely related plant taxa can differ considerably in size and arrangement (Tang et al. 2008). Chromosomal rearrangements have also been hypothesized to be involved in causing sterility in plants, and hindering recombination and thus reducing gene flow (Rieseberg 2001).

Genomic changes can generally be classified into two categories: numerical and structural. Structural rearrangements of chromosomes are typically initiated by one or more DNA double-

strand breaks, followed by a misrepair of the broken fragments. Numerical aberrations are typically caused by faulty spindle attachment in meiosis or mitosis, leading to a reduced or increased chromosome numbers, although some also result from structural chromosomal rearrangements like fusions and fissions. For reasons of clarity and comprehensibility, main processes in these categories will be presented separately, even though numerical and structural variations are interrelated – collectively, they produce evolutionary change (Weiss-Schneeweiss & Schneeweiss 2013).

1.1.1. Structural aberrations

Changes affecting chromosome structure can result in balanced or unbalanced gametes. While the former implies no loss or gain of genetic material, the latter encompasses chromosomal duplications or deficiencies and usually has severe consequences for the organism's reproductive ability, as it frequently leads to inviable gametes or zygotes (Rieseberg 2001). Both balanced and unbalanced rearrangements are to some extent unbalanced for progeny. Unbalanced rearrangements include insertions, deletions and duplications, while inversion and translocation are balanced ones.

Insertion

Insertions follow double strand breaks as additional DNA sequences are built into the chromatid during the repair phase. Lysák and Schubert (2013) name "recombination with extrachromosomal circular DNA" as possible mechanisms.

Deletion

Terminal deletion is the loss of a chromosomal terminal fragment. The resulting open double strand break (DSB) is "healed" either by synthesis of new telomeres, or by using telomeric sequences from sister chromatids (Lysák & Schubert 2013). Interstitial deletion, which requires two DSBs in the DNA double helix, describes the loss of an inner chromosome fragment and is repaired by joining the resulting ends. Some authors refer to this type of deletion as deficiency (e.g. Stebbins 1971).

Duplication

In the case of a duplication event, one or several chromosome fragments are doubled. The most frequent kind of this rearrangement is a tandem duplication, in which the segments that are duplicated are placed adjacent to one another, and thus, the DNA sequence repeats itself (Lysák & Schubert 2013). This change in chromosome structure is particularly important from an evolutionary point of view, as it provides material which can be altered while, for instance, an extra copy of the same sequence remains intact. In insertional duplications, the copied regions remain in separate parts of the same or even of different chromosomes altogether (Griffiths et al. 1999). In diploid organisms, three copies of the same sequence will exist after a duplication event. In the meiotic prophase of duplication heterozygotes, a loop is created in the unpaired duplicated region (Lysák & Schubert 2013).

Inversion

Inversions are rearrangements leading to balanced gametes as the overall amount genetic material remains the same. Inversions involve the change of directionality of a chromosomal segment; thus, the gene order within this segment is reversed. This restructuring is the result of a DSB and a subsequent inverted fusion (ligation). In pericentric inversions, breakpoints appear in both arms of the chromosomes. The centromere is therefore included in the inversion event and the arm ratio of the chromosomes might change (Lysák & Schubert 2013). Paracentric inversions on the other hand do not involve the centromere but appear only within one chromosomal arm (Schubert 2007). Such rearrangements typically result in partially inviable gametes(Lysák & Schubert 2013; Stebbins 1971; Griffiths et al. 1999).

Translocation

The structural rearrangement during which a chromosomal segment is transferred into a non-homologous chromosome is called translocation. In a reciprocal translocation, an exchange of genetic information takes place *between* two non-homologous chromosomes, while the transfer only takes place *from one chromosome to another* in non-reciprocal translocations. Similar to the way a loop is formed in inversion heterozygotes, chromosomes in a translocation heterozygote form a cross-shaped tetravalent of four chromosomes involved in translocation at the pachytene

stage of meiotic prophase I (Lysák & Schubert 2013). At a subsequent stage of prophase I and first metaphase of meiosis, this cruciform arrangement is opened and forms a ring (Stebbins 1971). It then depends on the way the spindle apparatus attaches whether viable and balanced gametes can be created or if the gametes will be unbalanced and unfit to survive. The former is more likely in the case of alternate segregation, i.e. if "the translocation chromosomes are pulled to one pole and 'normal' chromosomes to the other pole", while the latter is usually the result of adjacent segregation, when "either both 'adjacent' non-homologous or homologous centromeres pass to the same pole" (Lysák & Schubert 2013). Translocations and inversions have effects on genetic linkage, as they change the distance between genes on chromosomes and creates new gene combinations due to the rearrangement thus influencing the possibilities of successful crossing over (Stebbins 1971).

1.1.2. Numerical aberrations

The rearrangements described above not only alter chromosome structure and their genic content, but can also lead to changes in chromosome number. Generally, numerical aberrations are categorized into three groups: dysploidy, aneuploidy and polyploidy. Ploidy refers to the number of full chromosome sets a cell possesses (haploid, single set of chromosomes; diploid, two sets of chromosomes; triploid, three sets of chromosomes, etc.).

Dysploidy describes instances in which chromosome numbers change due to certain types of translocations via fusions or fissions with no significant loss or gain of DNA (Luceño & Guerra 1996). Dysploidy is an important evolutionary mechanism for chromosome number change because it does not create genetic imbalance (Weiss-Schneeweiss & Schneeweiss 2013). Most commonly the mechanism leading to dysploidy is the Robertsonian translocation or centric fusion, in which the DSBs occur in centromeres of two acrocentric or telocentric chromosomes followed by a fusion of the centromeric regions to create a metacentric/submetacentric larger fusion chromosome and an acentric fragment. The smaller fragment is acentric and thus typically lost, and the chromosome number of the cell changes (n-1, descending dysploidy). Effectively only a small reduction in DNA amount has taken place. Centric fission, on the other hand, increases the chromosome number (n+1, ascending dysploidy), as it involves the breakage of larger a

metacentric chromosome into two telo- or acrocentric ones within the centromeric region followed by telomere healing (Lysák & Schubert 2013).

Aneuploidy, on the other hand, creates a genetic imbalance as it involves duplication or deletion of one or two full chromosomes within a complement, thus a substantial number of genes is in higher copy number than required. Aneuploidy thus frequently leads to inviability or at least sterility of diploid organisms (Luceño & Guerra 1996). Consequences of aneuploidy are less severe if the anomaly occurs on small or gene-poor chromosomes, although the rate at which the change happens is distributed evenly across all chromosomes. The chromosome set of an aneuploid cell/organism differs from the regular chromosome complement in that it lacks one or two chromosomes or has an extra copy of one or two chromosomes (Weiss-Schneeweiss & Schneeweiss 2013; Griffiths 2000). Depending on the number of chromosomes that are gained or lost, the aberration is referred to as nullisomy ($2n$-2), monosomy ($2n$-1), trisomy ($2n$+1), tetrasomy ($2n$+2) etc. If, for instance, four copies each of two different chromosomes exist in a diploid organism, it is called *double* tetrasomic ($2n$ +2+2). The main cause of aneuploidy is chromosome nondisjunction due to improper spindle attachment in either of the two stages of meiosis, but also misdivision of chromatids during mitosis can lead to chromosome loss or gain. While in mammals, most types of aneuploidy are lethal, it is seen in individual diploid plant organisms but impairs progeny. In polyploid organisms (see below), aneuploidy is more tolerated but can be hard to distinguish from dysploidy (Weiss-Schneeweiss & Schneeweiss 2013).

1.1.3. Polyploidy

Polyploidy, also known as whole-genome duplication (WGD), is the process in which the entire chromosome set of an organism is multiplied and, as a result, the organism has three or more full copies of the haploid chromosome sets (Stebbins 1971; Wendel 2000; Leitch & Leitch 2008; Weiss-Schneeweiss & Schneeweiss 2013). Polyploidy stems from misdivisions in mitosis or meiosis, and frequently involves nondisjunction of whole chromosome sets (i.e. problems during spindle attachment), which leads to unreduced gametes (Leitch & Leitch 2008). Polyploidization events entail the gaining additional copies of complete genomes and is typically followed by genetic, epigenetic, and genomic structural changes aiming at cytological and genetic diploidization of the multiplied chromosome sets. Polyploidy in plants is often recurrent and

cyclical, which often leads to formation of polyploid series or polyploid complexes (Stebbins 1971; Leitch & Leitch 2008).

Typically, two different types of polyploids are recognized based on the mode of their origin: allo- and autopolyploids. Autopolyploids originate via multiplication of sets of chromosomes within one species or subspecies, which implies that the sets are identical or at least of very high similarity/homology (Otto 2007; Weiss-Schneeweiss et al. 2013). They are typically morphologically rather similar to their diploid progenitor, differing only in some characters, especially the size or morphology of their pollen grains, size of the plants and often experience low fertility due to the presence of more than two homologous chromosome sets, impairing bivalent pairing in meiosis (Stebbins 1971). Allopolyploids arise from interspecific hybridization of two species which are genetically different accompanied by genome duplication. These two types of polyploids differ in the way their chromosomes behave during meiosis (Ramsey & Schemske 1998). Polysomic inheritance and multivalent pairing, i.e. more than two homologous chromosomes pairing during meiosis I, are more common in autopolyploids than in allopolyploids, the latter showing more instances of bivalent formation, due to the presence of homoeologous genomes in allopolyploids (Otto 2007; Stebbins 1971), and disomic segregation. Exceptions, however, exist in both cases (Stebbins 1971; Weiss-Schneeweiss & Schneeweiss 2013). Allotetraploids formed by hybridization of two closely related species will more likely share high levels of synteny.

The effects of polyploidization on an organism are extensive, complex and still not fully understood. For instance, it is not entirely clear how a polyploid organism reverts to regular bivalent pairing mechanisms, although the process may be controlled genetically by "pairing genes" in some organisms (Leitch & Leitch 2008). Changes associated with polyploidy happen at the genomic, epigenomic and proteomic level, including "(retro)transposon mobility, sequence rearrangements and losses, gene silencing, DNA methylation changes, and chromatin remodeling" (Leitch & Leitch 2008). Over time, the large genomes of polyploids often experience downsizing and are reduced in size, accompanying diploidization processes (Leitch & Bennett 2004). Several evolutionary pathways are relevant for genes duplicated via polyploidy: some genes will experience functional diversification (neo- or subfunctionalization), some will be inactivated or lost, and some will retain their functions (Wendel 2000). Such diversification potential of the genes

therefore can be advantageous at particular levels of development and evolution. Polyploidy is thus considered to offer the raw material for the evolution to work on, allowing for creating more complex gene networks and leading to "biochemical and physiological flexibility, enhanced environmental adaptability, or the evolution of novel physiologies or morphologies" (Wendel 2000). Shortly after their formation, polyploids experience a revolutionary phase when the coexisting genomes experience various structural, genetic and epigenetic changes. This stage is usually the most important part of the process of diploidization: secondary reversal to diploid-like state. Once polyploids are able to pair in a diploid-like manner (bivalent formation; cytological diploidization) in meiosis and gene redundancy is reduced, a long evolutionary stage begins during which a subsequent round of polyploidization might occur.

Polyploidy as a process defines the evolutionary history of all eukaryotes. It is considered a key factor in the early diversification and speciation of animals, but appears to be even more prominent and ongoing in plants. While WGD is less significant in liverworts, gymnosperms and hornworts, it is common in green algae, lycopods and ferns (Weiss-Schneeweiss et al. 2013) and certainly one of the most important mechanisms in angiosperm evolution (Weiss-Schneeweiss & Schneeweiss 2013; Weiss-Schneeweiss et al. 2013; Wendel 2002). The frequency of polyploidy in angiosperms, i.e. the proportion of taxa that have experienced at least one episode of WGD, is estimated to lie somewhere between 30% to up to 70%, with most sources leaning towards 50% (Soltis et al. 2015). As diploid-like behavior is often restored following polyploidization, it can be difficult to establish how many cycles of WGD an angiosperm genome has experienced, and virtually all angiosperms are considered to be 'paleopolyploids', implying that their ancestral lineage has undergone ancient genome doubling (Wendel 2000; Soltis et al. 2009; Husband et al. 2013). Polyploidy is therefore seen as the key factor in angiosperm diversification (Jiao et al. 2011).

While polyploidy, especially autopolyploidy, was considered to be disadvantageous for higher plants in the past (Stebbins 1971), more recent studies recognize the process as crucial to the evolutionary success of angiosperms (Jiao et al. 2011). Polyploidy is a frequent and successful process contributing to the evolution of flowering plants, in contrast to aneuploidy. The latter is often highly problematic because it disrupts genetic balance, while polyploidy preserves it ("balance hypothesis", Otto 2007). Indeed, genes doubled by polyploidy tend to persist longer than those duplicated on a local genomic scale (Lynch 2007). Although it is not fully understood how

polyploidy specifically contributes to evolutionary success of lineages or organisms, it is widely recognized as a primary process in plant speciation, as well as the main cause of genomic redundancy.

1.2. Repetitive DNA

Single-or low copy coding sequences, regulatory DNAs, introns and promoters are important elements of the nuclear plant genomes, but their numbers are relatively similar across taxa. It is the repetitive DNA – sequence motifs which are present hundreds or thousand times in the genome – that constitutes the largest fraction of the nuclear genomes of plants and is responsible for the variation in genome sizes (Schmidt & Heslop-Harrison 1998; Leitch & Leitch 2008; Biscotti et al. 2015). There are several classes of repetitive DNAs, coding and non-coding, some of which with no known function. Tandem repeats contain copies of the same DNA sequence and include satellites, rDNAs and telomeres. Satellite DNAs, i.e. non-coding and tandemly repeated sequences, are often clustered around centromeres and telomeres. Telomeres are also a class of tandem repeats functioning as a physical cap at the ends of chromosomes that protects them, ensures linear replication and counteracting chromosome shortening due to the 'end-replication problem' (Schmidt & Heslop-Harrison 1998). For a comprehensive characterization of the 'end-replication problem' see, for example, Bryan (2009). Coding repetitive rRNA gene sequences are also tandem repeats at the 45S (=18S-5.8S-25S) rDNA loci, which also include various intergenic spacers and are associated with nucleolus during transcription (and thus referred to as nucleolus organizing regions; NOR) (Heslop-Harrison & Schwarzacher 2011; Weiss-Schneeweiss & Schneeweiss 2013), and 5S rDNA loci. These two markers are particularly useful for the identification of chromosomes due to very conserved sequences, their variability in loci localization and number, as well as their possible involvement in species differentiation due to their mobility within the genome (Heslop-Harrison 2000; Weiss-Schneeweiss et al. 2008; Weiss-Schneeweiss & Schneeweiss 2013). In contrast to those tandem repeats, dispersed repeats are typically dispersed across the whole genomes. They encompass various classes of transposable elements (class I: retrotransposons; class II: DNA transposons), which are responsible for their own replication, movement and integration into the host plant genome, and can therefore create new copies and/or change position in the genome (Heslop-Harrison & Schwarzacher 2011).

While repetitive DNA sequences have for a long time been deemed 'selfish elements' or even 'junk DNA' and largely disregarded in genomic studies (Mehrotra & Goyal 2014), their analysis is now considered critical for understanding plant genome composition and evolution. Relatively recent advances in methodologies, such as fluorescent *in situ* hybridization and next-generation sequencing (NGS; also known as high-throughput sequencing), allows more detailed chromosome mapping and characterization of repetitive DNA fraction, not limited to selected model organisms (Weiss-Schneeweiss & Schneeweiss 2013; Egan et al. 2012; Deschamps & Campbell 2010). New sequencing techniques permit rapid and cost-effective analyses of repetitive DNA, and enable researchers to address questions concerning genome variation, evolution and speciation at the level not attainable before.

1.3. Domestication as a driving force in evolution

Plant genome evolution is a result of various interrelated processes leading to morphological, physiological and genetic shifts in the affected species. One of these driving forces promoting chromosomal change, including but not limited to structural and numerical rearrangements, is domestication. Crop plants are frequently used as study subjects to analyze the mechanisms involved in genome evolution because of their young evolution and availability of archaeological and historical data related to their development and diversification (Meyer & Purugganan 2013). What is more, the basic mechanisms acting in the evolution of crop plants are the same as those that lead to genetic and phenotypic diversification in wild plants. Thus, the studies of domesticated model organisms allow for much broader inferences to be made, also for wild plants (Pickersgill & Heiser 1976).

Plant domestication essentially describes the selection of plants with desirable traits and the subsequent gradual accumulation of genetic, morphological and physiological divergences which distinguish domesticated taxa from their wild progenitors. Most researchers date the emergence of domestication around 13,000-10,000 years ago, and by 4,000 years ago, the domestication of most major crop species known today was achieved (Doebley et al. 2006; Purugganan & Fuller 2009). It is commonly agreed that most domesticated plant taxa have experienced a dramatic reduction in genetic diversity due to a 'domestication bottleneck', i.e. a change of the allele frequency following a rapid decrease of population size (Wright et al. 2005; Tang et al. 2010; Innan & Kim 2004). Early

farmers did preselect individual plants from a wild population to start the domestication process and then continued cultivating new generations from the best seeds, which led to a rapid loss of diversity. The extent of this loss, however, depends on the length of the domestication period as well as on the original population size (Eyre-Walker et al. 1998).

The reduction in genetic diversity does not occur consistently across all genes of crop plant genomes, but is less noticeable if the genes are neutral, i.e. have no effect on the desired phenotype (Doebley et al. 2006). Other genes which do influence the favored phenotype are, in addition to the bottleneck, affected by artificial selection, and genetic variation on these loci is minimal or can even be completely removed (Wright et al. 2005). In the evolution of crop plants, most alleles that were selected were already available in the ancestral gene pool. Novel mutations occurred for fewer traits when these mutations affected important developmental pathways (Doebley et al. 2006). Genome sequences of crop plants are examined to determine changes in diversity, organization and structure, and to understand which underlying adaptive mechanisms function as responses to artificial selection (Kantar et al. 2017). These findings can aid the understanding of broader, natural evolutionary processes and contribute to the improvement of future breeding efforts. Again, the development of fast and cost-effective high-throughput sequencing techniques is a crucial step in the analysis of such crop genomes.

2. THE GENUS *CAPSICUM*

The genus *Capsicum* – frequently referred to, among many other names, as “pepper”, “cayenne”, “chilli” or “red pepper” (for a more detailed entymology see Basu & De 2003) – is a group of flowering plants within the Solanaceae family. The nightshade family includes several other economically important genera such as *Solanum* (tomato, potato, eggplant), *Nicotiana* (tobacco) and *Petunia* (Edmonds 2012). Solanaceae include 3000-4000 species within about 90 genera, and the family is characterized by a high morphological and ecological diversity.

Some nightshade species have been used as food sources, medicinal drugs and ornamental plants for thousands of years. In addition to their significance as crops, several species within this family have functioned as models in classical and molecular genetic and genomic research (Gebhardt

2016). *Capsicum*, in particular, is suitable for cytogenetic research. Although chile peppers are known for their morphological diversity, their chromosome number is remarkably stable, with basic chromosome numbers of either $x = 12$ or $x = 13$ (cf. Moscone et al. 2007). Their relatively large genome sizes (cf. Moscone et al. 2003) do not result from polyploidization events (i.e. whole-genome duplications), but from the accumulation of large amounts of repeats of non-coding DNA within the chromosomes.

The following sections will characterize the study objects of this thesis, by providing an overview of the origin, phylogeny, domestication, morphology and economic importance within the genus *Capsicum*.

2.1. Origin and domestication

The genus consists of about 35 species worldwide, five of which are, according to most authors, considered domesticated (Carrizo García et al. 2013). These will be presented in section 2.3.1 of this thesis. The origin of the species is not fully understood, and numerous studies attempt to disentangle their phylogenetic relationships and impact of domestication on various aspects of their evolution. Including multiple lines of evidence can lead to conflicting results, which is why most studies on the genus' origin are not in complete agreement with each other. What is more, these studies generally include the analysis of only part of the genus, which makes it difficult to directly answer the question where domestication of *Capsicum* began. Recent reconstruction of the ancestral areas of eleven clades within the genus using Bayesian MCMC analysis revealed that the diversification started in western-north-western South America, continuing over the Amazonian lowlands to central and south-eastern Brazil, returning to central and western South America and eventually reaching Central America (Carrizo García et al. 2016). This study suggested that speciation occurred rapidly resulting in the large number of species known today.

There is general consensus that more than one region has to be considered when discussing the domestication process of chile peppers (Aguilar-Meléndez et al. 2009; Pickersgill 2007; Perry et al. 2007). Chile peppers were already consumed in Mesoamerica around 7000 BCE (Basu & De 2003). The wild chili known as "Piquin" was presumably collected from wild plants, before cultivation and domestication began around 5000 BCE. Plants which did not shed their ripe fruits

immediately were likely selected for breeding, as well as those with larger, more pendant pods (DeWitt & Bosland 2009). This makes *Capsicum* one of the earliest crops of the Americas, along with maize (see Benz 2001) or potato (see Hardigan et al. 2017).

The geographical origin of *Capsicum* domestication also received some attention in the past. Research in the field of paleobiolinguistics, i.e. a scientific branch within historical linguistics, established the time and place of *Capsicum* gaining significance for Native American peoples (Brown et al. 2013). Occurrences of names for chile pepper in major proto-languages of the Americas mainly revealed Mesoamerica and South America as possible areas of independent domestication of *Capsicum*. Other authors, such as Carvalho et al. (2014), claim that the evolutionary history of domesticated *Capsicum* species began independently in at least three regions. Analyses of three single- or low-copy nuclear loci of multiple *C. annuum* accessions indicated the Yucatan Peninsula as an important region for the domestication of this species (Aguilar-Meléndez et al. 2009). However, combined analyses from species distribution modeling and paleobiolonguistics with microsatellite genetic and archaeobotanical data concluded that the domestication of *C. annuum* initially occurred in northeastern and central-east Mexico (Kraft et al. 2014). This assertion partly coincides with findings from Pickersgill (2007), who, based on karyotype variation, named Mesoamerica as the region of domestication *for C. annuum*. Moreover, the valley of Bolivia south of the Amazon watershed is proposed for *C. baccatum*, and the tropical lowlands east of the Andes for *C. chinense*. Further research suggests that *C. pubescens* was domesticated in the mid-elevation southern Andes of Peru and Bolivia (Eshbaugh 1993, cited in Perry et al. 2007), while the case of *C. frutescens* is less clear, as the Carribean, Central America or southwestern Amazonia are all considered possible places of its domestication (Perry et al. 2007; Brown et al. 2013). Further research will be necessary to answer open questions concerning the origin of the domesticated *Capsicum* species.

2.2. Economic importance and use

Capsicum is part of the Solanaceae family, which includes a number of economically important crops such as tomato, potato, eggplant and tobacco (DeWitt & Bosland 2009). *Capsicum*, which is sometimes divided into the pungent chili peppers and sweet bell peppers in colloquial language, was brought from the American continent to Europe around 1493, where it was introduced by

Spanish explorers (Boswell 1949, cited in Lippert et al. 1966). From Spain, the cultivation of the new crop spread across Europe. A few decades later, Portuguese explorers from Brazil introduced *Capsicum* to the Indian market, and, according to reports of that time, peppers even reached China before 1800 (Sturtevant 1885, cited in Lippert et al. 1966). Long before the crop came to Europe, however, *Capsicum* was an essential part of the diet of the native peoples of the Americas (Basu & De 2003). Sweet bell peppers were a common ingredient in many dishes of these peoples, but the pungent varieties were brought to Europe at first, as their use as spices was regarded more valuable than use of the non-pungent varieties as vegetables (Heiser & Smith 1953).

Today, *Capsicum* plants are used in a variety of different ways. Over time, they have become an integral part in people's diets worldwide and are known for their high vitamin C content, although its content varies depending on the way of preparation. Consumed as fresh vegetables, for instance, red bell peppers have the highest high vitamin C content, which ranges from around 150 to 180 mg per 100 g fresh weight (Heiser & Smith 1953). When the ripe fruit is dried and ground, it is processed into cayenne pepper, paprika or chili powder – seasonings highly prevalent in several different cuisines all over the world. Food and spices are certainly the most common uses of *Capsicum*, although medicinal application of the plant is also widespread. Not only in indigenous healing traditional medicine, but also in Western medicine *Capsicum* is utilized to treat gastro-intestinal infections, respiratory problems, muscle soreness, lack of appetite, or function as pain killers or as aphrodisiac. What is more, the smoke of chilies was even used in warfare in pre-Columbian America and up until the twentieth century, and pepper spray is still a frequently applied tool for personal defense (Basu & De 2003), as it has an effect on mammalian ocular tissue (Krishnatreyya et al. 2018). Understanding the processes which result in such inflammation is vital for future studies on the treatment of inflammation, irritation and other ocular diseases.

Most of the uses of peppers are attributable to a naturally occurring alkaloid which is produced in glands in the placental tissue within the pods and belongs to the chemical group of capsaicinoids. Capsaicin ($C_{18}H_{27}NO_3$ or trans-8-methyl-N-vanillyl-6-nonenamide), is the active ingredient in chile peppers, and is therefore responsible for their pungency (Sharma et al. 2013). Contrary to popular belief, the seeds and ovary wall of the fruit are non-pungent, and the level of pungency differs greatly between varieties, assumed to be controlled by modifying genes as well as the ratio of seed and pod wall to placental tissue (Heiser and Smith 1953). Naturally, it acts as a chemical

defense against herbivores and fungi, and is described as an odorless, colorless, hydrophobic crystalline powder. The irritation caused by capsaicin is due to its binding to the receptor TRPV1, or transient receptor potential vanilloid subfamily number 1 (Sharma et al. 2013), in the pain pathway. After their initial excitation, neurons affected by the irritating compound undergo a "defunctionalization" period, during which they are not responding to a large number of stimuli. This refractory period after stimulation by capsaicin is, according to Sharma et al. (2013), "unique among naturally occurring irritant compounds," and is the basis for its use as a natural painkiller. Despite its worldwide significance, the regulation of capsaicin is not fully understood (Kim et al. 2014).

2.3. Phylogeny of the genus

Table 1– Classification of *Capsicum*

Taxonomic rank	
Order	Solanales
Family	Solanaceae
Subfamily	Solanoideae
Tribe	Capsicae
Genus	*Capsicum*

The classification and taxonomy within the genus *Capsicum* is not fully agreed upon, and there is no consensus about the number of species within the genus, but it is agreed that the number is more than 35 (Carrizo García et al. 2013). For the purpose of general clarification, this section will provide an overview of the genus' phylogeny, and will present in detail the five species which represent the domesticated chili peppers.

Over 10.000 varieties of chile peppers exist worldwide (Dewitt and Bosland 2009). Difficulties of classification can be observed throughout history, as attempts to classify chile peppers date back to the 16th century, where botanical books name either 33 or 27 species. In his 1753 work '*Species plantarum'*, Linnaeus named *C. annuum* and *C. frutescens* as the only species of the genus, before

adding *C. baccatum* and *C. grossum* in 1767. By the 19th century, the list of species grew to over 90, before returning to Linnaeus' original proposition (DeWitt and Bosland 2009). Irish (1898, cited in Smith & Heiser 1951) placed all commercial varieties of chili peppers under *C. annuum*, and considered the rest as varieties of *C. frutescens*. This belief was held for a long time, until Smith and Heiser (1957) organized the list of domesticated *Capsicum* into five species: *C. annuum*, *C. baccatum* (also: *C. pendulum*), *C. frutescens*, *C. pubescens* and *C. chinense*.

The history of the genus' classification is this complex because scholars originally relied on fruit morphology to classify varieties of *Capsicum*, while floral anatomy and other morphological characteristics were largely neglected (Basu & De 2003). Today, most authors agree that *Capsicum* consists of 35-40 species all native to the New World, with numerous varieties and cultivars (Carrizo Garcia et al. 2013; Carvalho et al. 2014; Romero-da Cruz et al. 2017). Classified according to their level of domestication, 30 species are wild (or semiwild), while five taxa are considered domesticated (Moscone et al. 2007), and a vast number of cultivars with unique characteristics and pod types exist within each of the five domesticated species.

A large sample of 34 species of *Capsicum* taxa was analyzed to determine the phylogenetic relationships within the genus based on sequences of *matK* and *psbA-trnH* intergenetic spacer, two plastid markers, as well as *waxy,* a single-copy nuclear gene (Carrizo García et al. 2016). The study confirmed that *Capsicum*, despite its morphological heterogeneity, is indeed a monophyletic genus, which, together with its sister genus *Lycianthes,* belong to the Capsicae tribe. An informal grouping of *Capsicum* into eleven clades was proposed. All the domesticated species were placed in a large well-supported superclade, but the interspecific relationships in this clade were less supported. Some morphological features, especially fruit color and giant cells in the mesocarp were found to be more informative than others to determine taxonomic relationships.

2.3.1. The domesticated chile peppers

Capsicum annuum

Capsicum annuum var. *annuum* is the most common of the five domesticated species, and yields the largest number of cultivars which are typically classified according to their pods. In a

commercial context as well as in the gardens of hobby botanists, *C. annuum* is the principal species, and it tops the list of peppers grown in Hungary, India, Mexico, China, Korea and the East Indies (DeWitt & Bosland 2009). Morphologically, *C. annuum* var. *annuum* is highly variable, as plant growth and leave size candiffer considerably between cultivars. The flowers are usually white or, in a few cases, purple, and the stems can be either hairy (pubescent) or smooth (glabrous). The wild form of *C. annuum* is the wild pepper chiltepín or chile piquin, *C. annuum* var. *glabruisculum*. This "mother of all chile peppers" (DeWitt & Bosland 2009) has small, round pods which turn red upon ripening.

Although the cultivars are in part easily distinguishable by their morphology, genetically and chromosomally they can be more difficult to characterize. *C. annuum* cultivars are difficult to be characterized by conventional chromosome staining methods due to the uniformity of their chromosomes (Romero-Da Cruz et al. 2017). Understanding chromosomal variation, however, is crucial to unveiling processes involved genomic evolution as well as the evolution of *Capsicum* as a genus. Therefore, the researchers combined several methods, fluorescent chromosome banding, silver staining and fluorescent *in situ* hybridization, to analyze and interpret variations among cultivars.

Capsicum chinense

Capsicum chinense was named by the Dutch botanist Kikolaus von Jacquinomist in 1776, who believed that this taxon was originally domesticated in China, which proved to be unsubstantiated claim. The species is grown mainly in tropical regions, most commonly in the Caribbean. Their pods are, like those of *C. annuum*, highly diverse, and their flowers are typically white with purple anthers and filaments. The plants generally have multiple stems, which makes them appear more compact. *C. chinense* has a unique constriction between the calyx and pedicel of the flower, which makes it more easily distinguishable from the other domesticated *Capsicum* species. The hottest known chili peppers belong to *C. chinense*, but non-pungent varieties of the species exist as well (DeWitt & Bosland 2009). This species also often hybridizes with *C. annuum*.

Capsicum pubescens

Capsicum pubescens is the only domesticated species in the genus of which no wild form exists. It is believed to have been domesticated around 6000 years ago. Like *C. chinense*, *C. pubescens* also has a unique feature among the species: the leaves of its varieties are hairy and its seeds are black. What is more, plants of *C. pubescens* can grow up to 9 meters, and corollas are purple with large nectar producing glands, the so called nectaries (DeWitt & Bosland 2009).

Capsicum frutescens

In contrast to the other domesticated species, *Capsicum frutescens* is less diverse in pod size, shape and color. The flowers are greenish-white, without any visible spots, and anthers and filaments are of purple color. Plants are rather small and more compact, ranging between 0,3 and 1,2 meters. The species also includes hot and sweet varieties, but its best-known cultivar *Capsicum frutescens* var. *tabasco*, the main ingredient of the world-famous hot sauce (DeWitt & Bosland 2009).

Capsicum baccatum

Lastly, *Capsicum baccatum* is the youngest of the species, and was domesticated about 2500 years ago. *C. baccatum* exists in its wild form, *C. baccatum* var. *baccatum*, as well as in its domesticated form, *C. baccatum* var. *pendulum* and *C. baccatum* var. *umbilicatum*. It is most commonly grown in South America, and is morphologically highly diverse. Plants can be up to 1.5 meters high and, like *C. chinense*, have multiple stems with white flowers, which frequently show green, yellow or brown spots (DeWitt & Bosland 2009).

2.4. Morphology

Capsicum is renowned for its large variety of pods, the colorful berries of various shapes, sizes, and pungency levels, ranging from sweet bell peppers to hot jalapeños. Besides their ability to adapt to different climates, the multitude of different pod types is an important reason for the large economic interest in the genus. The berries are either erect or pendant, and change color according to their level of maturity. The color spectrum of ripe fruit includes red, yellow, orange, brown, green or white, while young, immature *Capsicum* pods are mainly green or white, but yellow or

purple varieties exist as well. The ripening process is complete after around 130 days, but fruit can already be consumed after only 70 days (DeWitt & Bosland 2009).

In contrast to its rich fruit diversity, the genus' floral characteristics are relatively little variable (Basu & De 2003). The flowers are relatively small (ca. 1.5 cm in diameter) and their orientation is typically pendant, the pedicels hanging down from the stem, with either a white or purple corolla and diverse pigmentation patterns. Within Solanaceae, *Capsicum* is exceptional because of its cup-shaped calyx, with 5-10 teeth functioning as nerve prolongations (Carrizo García et al. 2016). Before molecular methods were used to taxonomically characterize *Capsicum* plants, flower characteristics were therefore mainly considered in the classification process (e.g. Raghavan & Venkatasubban 1940). Smith and Heiser (1951), for example, mention pedicel number and length as well as petal color as useful characteristics to distinguish between *C. annuum* and *C. frutescens*. Like the pods, leaves also vary in size but are generally ovate with only one leaf per node, resulting in a spiral pattern (DeWitt & Bosland 2009).

In their native climate, *Capsicum* species are perennial, branched shrubs (Basu & De 2003). Typically facing colder climates, the domesticated members of the *Capsicum* genus are annuals (DeWitt & Bosland 2009). Shorter lifespan is however not the only characteristic distinguishing the domesticated *Capsicum* varieties from their wild progenitors. Pickersgill (2007) lists several morphological changes which occurred during the domestication process of *Capsicum* and conform to the familiar 'domestication syndrome' – a term frequently used in discussions about crop evolution. The term describes the co-occurrence of certain traits observed in domesticated plants which distinguish them from their wild ancestors (Harlan et al. 1973; Hammer 1984). These typically include larger and more robust plants, fewer mechanisms of seed dispersal, as well as higher grain size (Gross and Olsen 2010). Generally, the part that is harvested shows the most drastic changes, which is also the case for *Capsicum*. The pods of wild *Capsicum* plants are red, berry-like, small fruits (often referred to as "bird peppers"), and are shed from the plant upon reaching maturity (Aguilar-Meléndez et al. 2009). These characteristics are typical for seed dispersal by birds (ornithochory), as the color is attractive to birds and they can easily remove the fruit from the calyx (Carvalho et al. 2014). What is more, birds seem to be immune to the capsaicinoids, which function as protection from other mammals, insects or fungi (DeWitt & Bosland 2009).

Domesticated *Capsicum* do not disperse easily of their fruit, but the pods stay attached to the plant as long as possible. As the development of an abscission zone was lost during the domestication process, cells attaching the pods to the stem cannot separate from the surrounding tissue, and the fruit remain firmly attached. What is more, the pods, seeds and flowers are much larger than those of wild *Capsicum* plants, although, interestingly, not resulting from an increase in chromosome number (Pickersgill 2007). Although domesticated plants are in many cases more robust and larger than their wild ancestors, this is not true for *Capsicum*. Plant morphology varies greatly, but the domesticated plants hardly exceed 1 m, while wild varieties generally reach 1 m and above (DeWitt & Bosland 2009). Wild *Capsicum* plants are self-pollinating with occasional cross-pollination by insects (DeWitt & Bosland 2009). However, another characteristic resulting from human selection and cultivation is the plants' increasing inability to survive without human intervention. Domesticated varieties, therefore, rely on humans for successful reproduction, as their loss of dispersal and extreme pod size (compared to the wild progenitors) impede reproductive measures.

The basic chromosome number in the majority of *Capsicum* species is $x = 12$ (Huskins & La-Cour 1930), although $x = 13$ is present in a few wild species (Moscone et al. 1995). A reduction of genetic diversity has been reported in domesticates of *Capsicum* in comparison to their wild progenitors. Aguilar-Meléndez et al. (2009) analyzed nucleotide diversity at three nuclear loci, *Dhn*, *G3pdh*, and *Waxy*, in *C. annuum* var. *annuum* L. and semiwild *C. annuum* var. *glabriusculum*. A decrease in genetic diversity of 10-17% has been inferred, although higher numbers were expected based on studies of other crop species such as maize. The reason has been suggested to be resulting from sampling scheme, as semiwild populations are genetically less different from domesticated varieties than pure wild ones. This has been supported by another study (González-Jara et al. 2011), which reported a reduction of genetic variation of up to 40 % in cultivated *C. annuum* var. *glabriusculum* in comparison to wild populations of the pepper more commonly known as "chiltepin." These data further emphasize the importance of the conservation of wild crop relatives, as a source of genetic variation for cultivated taxa.

3. AIMS OF THE PRESENT STUDY

Economically, *Capsicum* has been considered a genus of great relevance for a long time. To this day, however, chile peppers have been the subject of far fewer studies that other crop plants such as maize, or even other members of the Solanaceae family like tomato. The present study aims to provide a cytogenetic characterization of 24 cultivated accessions belonging to either *C. annuum* or *C. chinense*, and to interpret the collected data in an evolutionary context. From each accession, nuclear DNA content as well as karyotypes were established. Interspecific and intraspecific morphological variations and similarities of karyotypes were examined, and a more detailed insight into genome evolution was gained by mapping rRNA genes in selected accessions using fluorescent *in situ* hybridization to further support the karyotypic analysis. Additionally, this study attempts to identify and characterize repetitive DNA fractions within the genomes of the two *Capsicum* species using next-generation sequencing (NGS) technology. The study provides insights into the genomic diversity of an economically significant plant group. It will aid the cytogenetic characterization of *Capsicum*, and can be instrumental to connecting genomic, morphological and phylogenetic information which has so far been generated in the genus. Moreover, it encourages the understanding of evolutionary processes, specifically those influenced by domestication.

PART II – Analyses of Selected Cultivated Chili Peppers

4. MATERIALS AND METHODS

4.1. Plant material

The plant material studied is presented in Table 2. We selected 12 cultivars/varieties of *C. annuum* and *C. chinense* (cultivated by Stekovics GmbH) and labelled them 1C-12C1. 15 individuals of the species *C. chinense* and 9 individuals of *C. annuum* were characterized. From each variety, two plants were grown in the greenhouse, except for 11C (three plants) and 12C1 (one plant) resulting in a total number of 24 plants.

Table 2 – List of cultivars

Label	Variety	Species
1C	Billy Goat	*C. chinense*
1C1	Billy Goat	*C. chinense*
2C	Trinidad Moruga Yellow	*C. chinense*
2C1	Trinidad Moruga Yellow	*C. chinense*
3C	Paradeispaprika Lozi	*C. annuum*
3C1	Paradeispaprika Lozi	*C. annuum*
4C	Jelly Bean	*C. chinense*
4C1	Jelly Bean	*C. chinense*
5C	Japonaise	*C. chinense*
5C1	Japonaise	*C. chinense*
6C	Beni Highland	*C. chinense*
6C1	Beni Highland	*C. chinense*
7C	Lippenstift	*C. chinense*
7C1	Lippenstift	*C. chinense*
8C	Large Orange Thai	*C. chinense*
8C1	Large Orange Thai	*C. chinense*
9C	Gänseschnabel	*C. annuum*
9C1	Gänseschnabel	*C. annuum*
10C	Apfelpaprika	*C. annuum*
10C2	Apfelpaprika	*C. annuum*
11C	Jalapeno	*C. annuum*
11C1	Jalapeno	*C. annuum*
11C2	Jalapeno	*C. annuum*
12C1	Habanero	*C. chinense*

Fig. 1 – Fruits of the 12 analyzed *Capsicum* varieties; (A) *C. chinense* 'Billy Goat'; (B) *C. chinense* 'Trinidad Moruga Yellow'; (C) *C. annuum* 'Paradeispaprika Lozi'; (D) *C. chinense* 'Jelly Bean'; (E) *C. chinense* 'Japonaise'; (F) *C. chinense* 'Beni Highland'; (G) *C. chinense* 'Lippenstift'; (H) *C. chinense* 'Large Orange Thai'; (I) *C. annuum* 'Gänseschnabel'; (J) *C. annuum* 'Apfelpaprika'; (K) *C. annuum* 'Jalapeno'; (L) *C. chinense* 'Habanero' (Photos: Sophie Rothen)

4.2. Flow cytometry

Isolation of the nuclei and propidium iodide staining

The protocol from Temsch (2003), which follows in principle Baranyi and Greilhuber (1996) was used to prepare the plant material for flow cytometry (FCM). To measure the DNA content, the plant cells were stained with propidium iodide (PI), a fluorescent dye which binds to DNA (Ormerod 2009). *Solanum pseudocapsicum* was used as the internal standard to provide a reference point for the measurement.

One or two young leaves (or about 25mg leaf material) were collected from every *Capsicum* individual analyzed, as well as from the standard species. The leaves of each individual were chopped together with the standard organism with a razor blade inside a petri dish with 1.1 ml isolation buffer (Temsch 2003), following the chopping method outlined by Gailbrath et al. (1983). Each round of preparations included one preparation with only material from the standard as a reference point for the measurements. Isolated nuclei suspensions were filtered through a 30µm nylon mesh into marked test tubes to remove the debris consisting of mainly cell walls. 50µl RNaseA (concentration: 0.15mg RNase A per ml suspension of nuclei; Temsch 2003) was added to each preparation and incubated in a water bath at 37°C for 30 minutes to allow for the digestion of the RNAs. Subsequently, each preparation was divided into two tubes. 4ml PI solution (concentration: 1mg PI per ml A. dest; Temsch 2003) was added to one of the two tubes before they were incubated in the refrigerator for 20 to 60 minutes. The second tube was kept in the refrigerator as backup.

Measurement procedure

Following the staining, measurement of the cellular DNA content was performed by flow cytometry. This process is generally used as a form of analysis of microscopic particles in suspensions, mainly for cell cycle analysis or DNA measurement, to determine the ploidy level of the cell material (Nunez 2001). In the flow cytometer, a fluorescent signal is being emitted from the stained material, which is in turn measured and analyzed by a detector. These signals are transformed into electric ones, and is depicted in a histogram, which represents the relative DNA content of an organism. The standard is necessary as a reference point to establish the absolute

DNA content. In this case, *Solanum pseudocapsicum* was used, with a DNA of 1C = 1.2946 pg. S*olanum pseudocapsicum* is less commonly used as a standard, but nevertheless useful due to its availability and low CVs (Temsch et al. 2010).

The measurements were performed using a CyFlow ML flow cytometer (Partec, Münster, Germany; green laser equipped). Each round of measurements was commenced by measuring the standard alone. Then, each sample was measured three times, each with a total "count" number of 3330 nuclei detected by the cytometer. FloMax software (Partec) was used to collect and analyze the data obtained from the cytometer. The coefficient of variation (CV) of G_0/G_1 peaks calculated from the standard deviation (Ormerod 2009) typically stayed below 3%. However, in a few cases a higher CV was deemed acceptable as well (between 5 and 10%). A more detailed explanation on how the data was interpreted will be given in chapters 5 and 6 (Results and Discussion). The 1C-value of a sample was calculated by dividing the mean of the G_1-peak ratios of the sample by the mean of the G_1-peak ratios of the standard, and multiplying the result by the 1C-value of the standard:

1C-value of sample in pg = [(mean G_1-peak ratio of sample) / (mean G_1-peak ratio of standard)] * standard 1C-value in pg

4.3. Chromosome counts

Chromosome pre-treatment

Healthy root tip meristems were collected from the pot-grown plants in the greenhouse and placed in a small bottle of water. In the lab the roots were then cleaned of soil and transferred into 0.002 M 8-hydroxyquinoline for 2.5 hours at room temperature and another 2.5 hours at 4°C.

Fixation

The 8-hydroxyquinoline solution was carefully removed with a pipette, and the material was washed twice with tap water to ensure that no solution was left on the root tips. Immediately afterwards the material was fixed in freshly prepared fixative (3 ethanol: 1 acetic acid), and was stored in darkness at room temperature for up to 12 hours. The material was then transferred to -20°C and stored for several months until use.

Feulgen staining

In order to examine the chromosomes under the microscope, the root tips were stained using the Feulgen technique (Fukui & Nakayama 1996). The root tips were removed from the fixative, draining the excess liquid in the process. Approximately three root tips were transferred into a tube filled with 5N HCl (VWR, Austria) and incubated for 20 minutes at room temperature. Subsequently, the HCl was removed from the tube and the material was washed twice with tap water. The excess water was drained and replaced with Schiff's reagent (Sigma, Austria), and the material was incubated at room temperature in darkness for 60 minutes to ensure proper staining. The stained material was placed in a drop of 60% acetic acid onto a pre-cleaned glass slide. The meristems were dissected under the dissecting microscope using entomological needles, while the other tissue material was removed. A coverslip was carefully placed on the slide and the material was spread out by lightly tapping with a needle. Lastly, the material was squashed to spread and flatten the chromosomes by placing a piece of filter paper over the coverslip and gently pressing down, making sure the coverslip does not move in the process. The preparations were then examined under an Axio-Imager A1 microscope (Carl Zeiss, Vienna, Austria) in order to find mitotic metaphase plates and establish the chromosome numbers of the cells. Chromosome spreads were then photographed using a CCD camera and software AxioVision 4.8 (Carl Zeiss, Vienna, Austria).

4.4. Fluorescent *in situ* hybridization

Fluorescent *in situ* hybridization (FISH) was performed to map 45S and 5S rRNA gene loci in three of the 24 individuals, *C chinense* 'Japonaise' (5C1), *C. chinense* 'Billy Goat' (1C1) and *C. annuum* 'Apfelpaprika' (10C).

Chromosome preparations

Chromosomes were prepared by enzymatic digestion and squashing as described in Jang and Weiss-Schneeweiss (2015) with minor modifications. After two washes in citric buffer (pH 4.8), the root meristems were incubated in the enzyme mixture (0.4% pectolyase, 0.4% cytohelicase, 1% cellulase in citric buffer, pH 4.8) for 18 minutes at 37°C. The enzyme mixture was removed, the

tissue washed with citric buffer and the material was transferred to a pre-cleaned slide and squashed in a drop of 45% acetic acid. The quality of the preparation was checked under phase-contrast before the cover slip was removed on dry ice. Preparations with about 15 well spread metaphases were stored at 4°C and used for FISH.

Fluorescence in situ *hybridization*

FISH was carried out as described in Jang and Weiss-Schneeweiss (2015). Refixation of the preparations was done in ethanol : acetic acid (3 : 1), and the slides were subsequently pre-treated with RNase A (100 µg/ml; 60 min at 37°C) and pepsin (10mg/1ml; 20 min at 37°C). Subsequently, a hybridization mix containing 10% dextran sulphate, salmon sperm blocking DNA (100ng/µl), 0.02 x SSC and 50-100ng of each of the two labeled probes was prepared. The probes used for FISH were 5S rDNA (genic region of 120 bp isolated from *Melampodium montanum*) in plasmid pGEM-T Easy, labelled directly by PCR with biotin (Roche, Vienna, Austria) and 18S rDNA (part of 45S rDNA array isolated from *Arabidopsis thaliana*) in a plasmid pSK+ labeled using a digoxygenin-nick-translation kit according to the manufacturer's instructions (Roche, Vienna, Austria). The hybridization mixture was denatured at 95°C for 2-5 minutes and immediately after cooled on ice for 5 minutes and applied to the preparations. Preparations with the denatured hybridization mix (8µl/slide) were denatured together on a PCR *in situ* block at 72°C for 4 minutes before incubation of the slides overnight at 37°C. Stringent washes were performed in 2 x SSC at room temperature for 10 minutes, followed by washes in 2xSSC, 0.1xSSC and 2xSSC + 0.2% Tween20 at 39°C for 5 minutes each. To detect the labeled probes, slides were incubated in blocking solution (1% BSA in 2xSSC + 0.2% Tween20) for 20 minutes at 37°C, followed by incubation with antibody solution (anti-digoxygenin conjugated with FITC to detect digoxygenin, and streptadivin conjugated with Cy3 to detect biotin) for 60 minutes at 37°C. Slides were washed twice in 2xSSC for 7 minutes at 42°C, and once in 2xSSC + 0.2% Tween20 for another 7 minutes in 42°C. The slides were mounted in antifade buffer (Vectashield containing DAPI) and stored in 4°C overnight. Preparations were analyzed with an Axiolmager M2 epifluorescent microscope (Zeiss, Vienna, Austria). Images were captured with a CCD camera and processed using AxioVision 4.8 (Zeiss, Vienna, Austria) with only those functions that apply to all pixels of the image equally.

4.5. Next-generation sequencing

Genomic DNA isolation, library preparation and sequencing

Genomic DNAs (gDNAs) of two analsed individuals (*C. annuum* accession C213 and *C. chinense* accession C14; cultivated in Botanic Garden of The University of Vienna) were isolated using a modified CTAB protocol (Doyle and Doyle 1987; Jang and Weiss-Schneeweiss 2015) and checked for quality on an agarose gel and for concentration using a fluorospectrophotometer and Quant-iT Picogreen dsDNA assay kit (PeqLab). Two independent libraries prepared for each individual were sequenced separately. Fragmentation resulting in 600–800 nt DNA fragments and library preparation were performed at the sequencing facility of CSF-NGS (Vienna Biocenter, Austria). The samples were sequenced using the 150 nt paired-end technology on an Illumina HiSeq2500 machine (Illumina, San Diego, CA, USA).

Analysis of the repetitive fraction of the genome using RepeatExplorer

Read pairs were filtered for quality. Adaptors and reads derived from plastid and mitochondrial genomes as well as the *PhiX* spike-in DNA (Illumina) were removed from the datasets. The reads were trimmed to 140 nt and assigned a unique three-letter species identifier for use in comparative clustering.

The reads were analyzed using the RepeatExplorer pipeline (Novák et al., 2010; 2013). Clustering analyses were first performed individually for each of the two species using the maximum number of reads, but the reads of two independently sequenced libraries were pooled together to detect biases due to library preparation. The reads were analyzed using the default settings in Repeat Explorer (Novák et al. 2013).

Briefly, an all-to-all blast comparison of all reads was used to generate a graph where the reads are connected by edges reflecting their similarity. The resulting clusters represent different repeat families (Novák et al. 2010). The annotation of the clusters was performed automatically (implemented in the RepeatExplorer pipeline) and verified manually. This allowed determining the type of elements making up the individual clusters and the genome as a whole.

5. RESULTS

5.1. Flow cytometry

The genome sizes of all 24 individuals were measured using FCM. The results are presented in Table 3 and include the mean 1C-values as pg of DNA, the coefficient of variation (CV) and standard deviation (S.D.). The number of runs was three per individual. The term "1C-value", measured in DNA picogram (pg), denotes the DNA content of an unreplicated, haploid nucleus with the chromosome number *n* (Greilhuber & Doležel 2009).

The following screenshot (Fig. 2) shows how the data is presented by the program FloMax (Partec). The lower part of the image shows the settings. In the first screenshot (data of the individual 5C), the speed is 0.1 and the FL1 wavelength is set at 323. In the resulting histogram, three peaks of the measured fluorescence intensity are visible on a scale from 0 to 1000. The G1-peak of the standard *Solanum pseudocapsicum* is set close to 100 with each measurement. The peaks are, from left to right, G1-peak and G2-peak of *Solanum pseudocapsicum* at 100.35 and 203.31 respectively, and the G1-peak of individual 5C (*C. chinense* 'Japonaise') at 266.7. The screenshot below shows the results of individual 5C1, with G1-peak and G2-peak of the standard at 107.43 and 216.48 respectively, and the G1-peak of 5C1 at 432.61.

5.2. Chromosome counts

On the subsequent pages, the results of the chromosome counts following Feulgen staining are presented. For each individual, a photo of a mitotic metaphase plate as well as a corresponding karyogram and idiogram are given. Black dots were added to the idiograms when satellites were clearly visible. For 23 of the 24 individuals that were examined, a chromosome number of $2n = 2x = 24$ was found. One individual of *C. chinense*, 5C1, showed a higher chromosome number, $2n = 3x = 36$. A more detailed discussion of the findings will be provided in section 6.2. Possible inconsistencies, e.g. discrepancies of chromosome sizes in idiograms, are most likely caused by subjective choices made in arranging the chromosomes and by the subsequent measuring procedure.

Table 3 – Genome sizes (expressed as 1C value; pg) and total haploid chromosome lengths (HCL) of all analyzed *Capsicum* accessions

Accession	mean 1C-value (DNA pg)	CV% of measurements	S.D. (pg)	HCL (µm)
1C (*C. chinense* 'Billy Goat')	3.4403	0.1668	0.0057	42.4
1C1 (*C. chinense* 'Billy Goat')	3.5100	0.1129	0.0040	38.4
2C (*C. chinense* 'Trinidad Moruga Yellow')	3.5035	0.1018	0.0036	36.1
2C1 (*C. chinense* 'Trinidad Moruga Yellow')	3.4947	0.0851	0.0030	45.1
3C (*C. annuum* 'Paradeispaprika Lozi')	3.4414	0.1425	0.0049	50.0
3C1(*C. annuum* 'Paradeispaprika Lozi')	3.5796	0.4745	0.0170	34.3
4C (*C. chinense* 'Jelly Bean')	3.4936	0.0056	0.1609	42.0
4C1 (*C. chinense* 'Jelly Bean')	3.5044	0.1531	0.0054	41.4
5C (*C. chinense* 'Japonaise')	3.4367	0.1763	0.0061	42.7
5C1 (*C. chinense* 'Japonaise')	5.2253	0.1349	0.0070	37.3
6C (*C. chinense* 'Beni Highland')	3.4506	0.1961	0.0068	34.4
6C1 (*C. chinense* 'Beni Highland')	3.5151	0.2510	0.0088	37.6
7C (*C. chinense* 'Lippenstift')	3.4707	0.0697	0.0024	48.8
7C1 (*C. chinense* 'Lippenstift')	3.4676	0.3693	0.0128	43.2
8C (*C. chinense* 'Large Orange Thai')	3.4536	0.2302	0.0080	46.6
8C1 (*C. chinense* 'Large Orange Thai')	3.4773	0.0536	0.0019	41.8
9C (*C. annuum* 'Gänseschnabel')	3.4283	0.2432	0.0083	31.0
9C1 (*C. annuum* 'Gänseschnabel')	3.4081	0.1634	0.0056	36.5
10C (*C. annuum* 'Apfelpaprika')	3.4162	0.1395	0.0048	42.2
10C2 (*C. annuum* 'Apfelpaprika')	3.4283	0.0998	0.0034	44.5
11C (*C. annuum* 'Jalapeno')	3.4249	0.3676	0.0126	42.3
11C1 (*C. annuum* 'Jalapeno')	3.4622	0.3446	0.0119	38.9
11C2 (*C. annuum* 'Jalapeno')	3.4094	0.1519	0.0052	50.0
12C1 (*C. chinense* 'Habanero')	3.4449	0.1903	0.0066	33.4

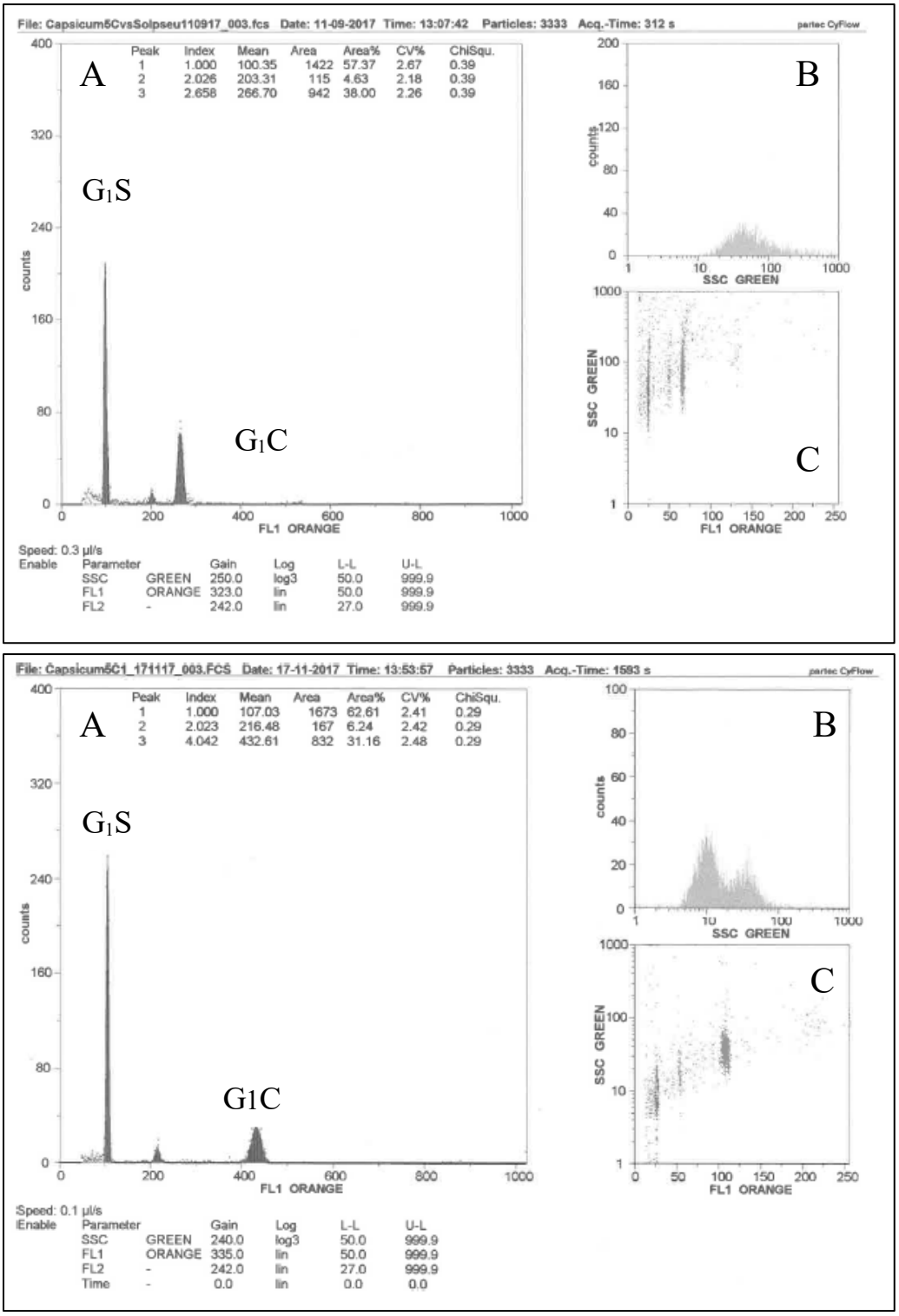

Fig. 2 – Examples of flow cytometry histograms of genome size measurements of *Capsicum chinense* and *C. annuum*. (A) Histogram of the relative DNA content (Individual 5C, Standard: *Solanum pesudocapsicum*, 3rd run); (B) side scatter histogram; (C) cytogram fluorescence (FL1) vs side scatter.

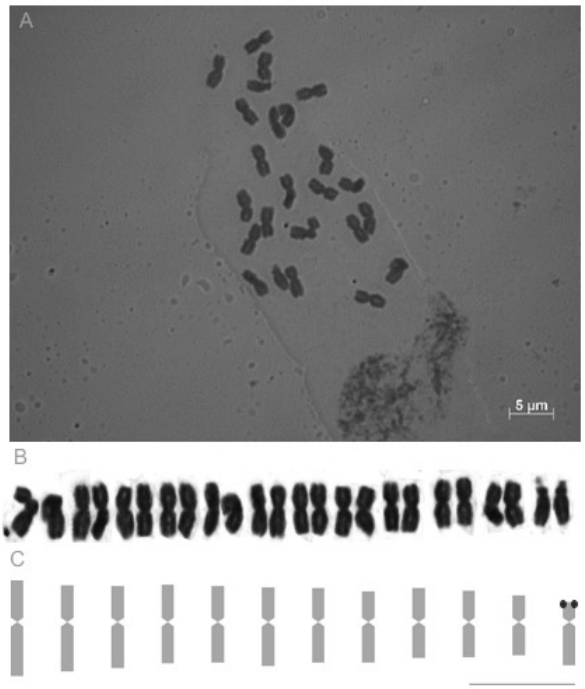

Fig. 3 – *C. chinense* variety 'Billy Goat', accession 1C: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram with satellites (NORs) indicated as black full circles; scalebar, 5 µm.

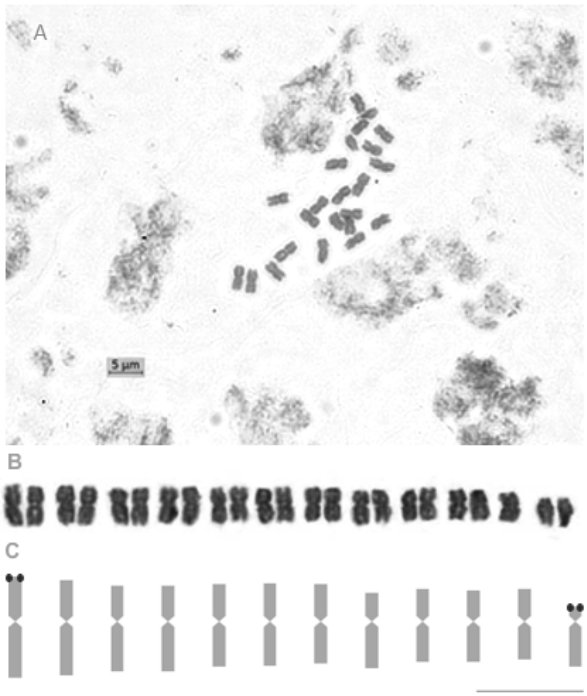

Fig. 4 – *C. chinense* variety 'Billy Goat', accession 1C1: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram with satellites (NORs) indicated as black full circles; scalebar, 5 µm.

The chromosome number for accession 1C, *C. chinense* 'Billy Goat', was $2n = 2x = 24$ (Fig. 3, A). Accession 1C showed 11 metacentric or submetacentric chromosome pairs and one acrocentric chromosome pair, possibly carrying a satellite (Fig. 3, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 42.4 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 3, C. The genome size (1C value) of accession 1C was 3.4403 pg (Table 3).

The chromosome number for accession 1C1, *C. chinense* 'Billy Goat', was $2n = 2x = 24$ (Fig. 4, A). Accession 1C1 showed 11 metacentric or submetacentric chromosome pairs, the largest of which carried a satellite, and one acrocentric chromosome pair, possibly carrying a satellite (Fig. 4, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 38.4 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 4, C. The genome size (1C value) of accession 1C was 3.5100 pg (Table 3).

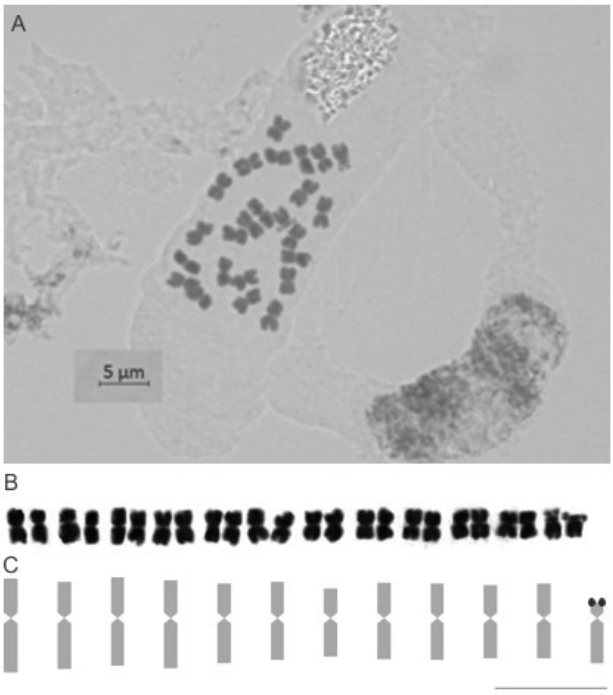

Fig. 5 – *C. chinense* variety 'Trinidad Moruga Yellow', accession 2C: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram with satellites (NORs) indicated as black full circles; scalebar, 5 µm.

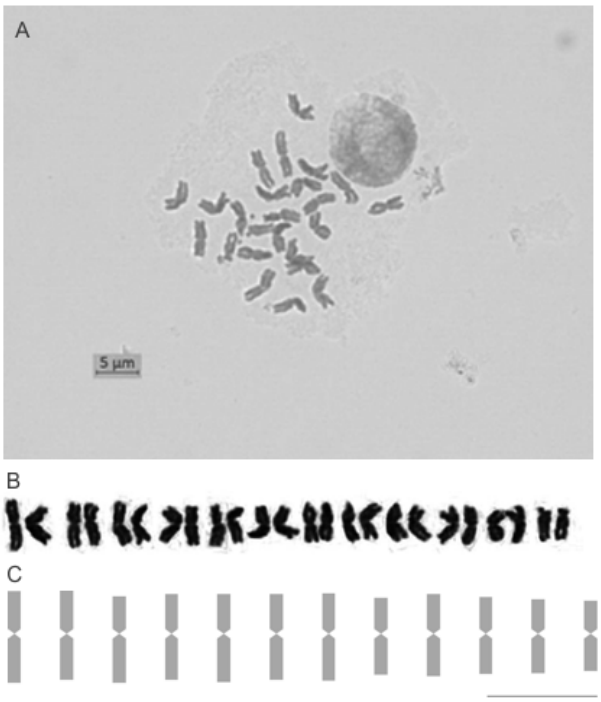

Fig. 6 – *C. chinense* variety 'Trinidad Moruga Yellow', accession 2C1: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram; scalebar, 5 µm.

Where you have satellites indicated, you have to mention it in the legend.

The chromosome number for accession 2C, *C. chinense* 'Trinidad Moruga Yellow', was $2n = 2x = 24$ (Fig. 5, A). Accession 2C showed 11 metacentric or submetacentric chromosome pairs and one acrocentric chromosome pair carrying a satellite (Fig. 5, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 36.1 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 5, C. The genome size (1C value) of accession 2C was 3.5035 pg (Table 3).

The chromosome number for accession 2C1, *C. chinense* 'Trinidad Moruga Yellow', was $2n = 2x = 24$ (Fig. 6, A). Accession 2C1 showed 11 metacentric or submetacentric chromosome pairs, the largest of which possibly carrying a satellite, and one acrocentric chromosome pairs with a putative satellite (Fig. 6, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 45.1 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 6, C. The genome size (1C value) of accession 2C1 was 3.4947 pg (Table 3).

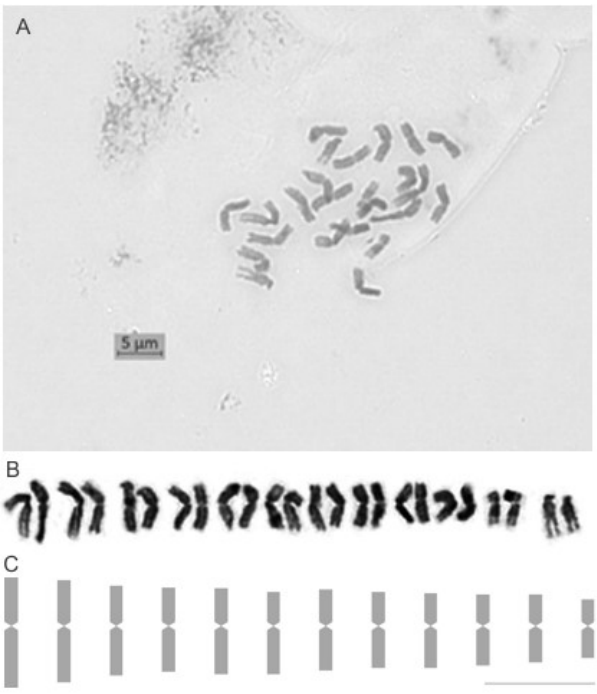

Fig. 7 – *C. annuum* variety 'Paradeispaprika Lozi', accession 3C: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram; scalebar, 5 µm..

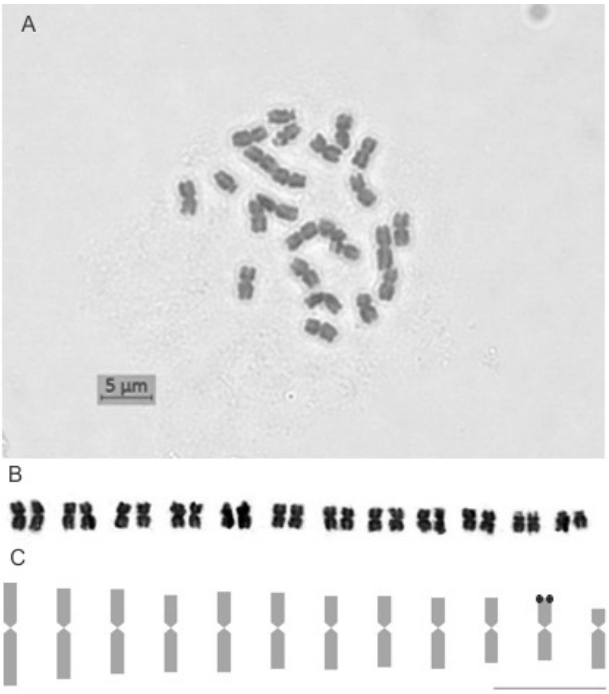

Fig. 8 – *C. annuum* variety 'Paradeispaprika Lozi', accession 3C1: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram with satellites (NORs) indicated as black full circles; scalebar, 5 µm.

The chromosome number for accession 3C, *C. annuum* 'Paradeispaprika Lozi', was $2n = 2x = 24$ (Fig. 7, A). Accession 3C showed 10 metacentric or submetacentric chromosome pairs and two acrocentric chromosome pair (Fig. 7, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 50.0 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 7, C. The genome size (1C value) of accession 3C was 3.4414 pg (Table 3).

The chromosome number for accession 3C1, *C. annuum* 'Paradeispaprika Lozi', was $2n = 2x = 24$ (Fig. 8, A). Accession 3C showed 10 metacentric or submetacentric chromosome pairs and two acrocentric chromosome pairs, one of them carrying a satellite (Fig. 8, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 34.3 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 8, C. The genome size (1C value) of accession 3C was 3.5796 pg (Table 3).

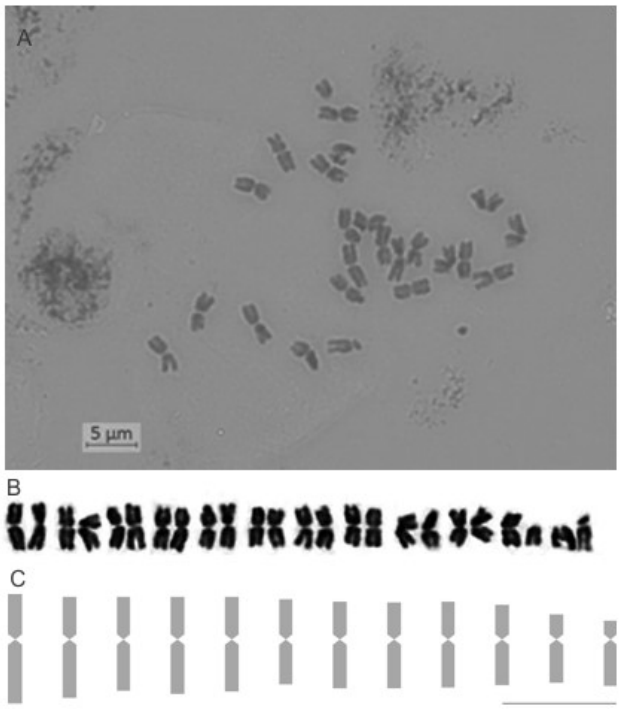

Fig. 9 – *C. chinense* variety 'Jelly Bean', accession 4C: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram; scalebar, 5 µm.

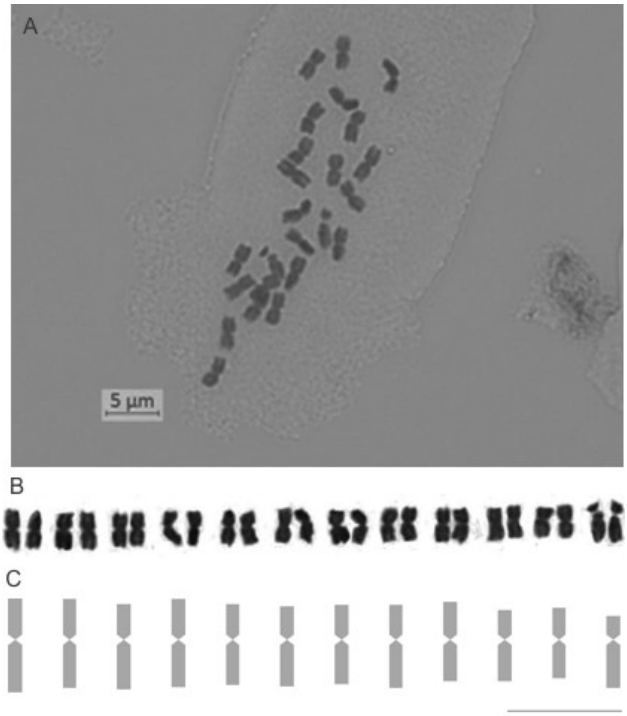

Fig. 10 – *C. chinense* variety 'Jelly Bean', accession 4C1: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram; scalebar, 5 µm.

The chromosome number for accession 4C, *C. chinense* 'Jelly Bean', was $2n = 2x = 24$ (Fig. 9, A). Accession 4C showed 11 metacentric or submetacentric chromosome pairs and one acrocentric chromosome pair (Fig. 9, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 42.0 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 9, C. The genome size (1C value) of accession 4C was 3.4936 pg (Table 3).

The chromosome number for accession 4C1, *C. chinense* 'Jelly Bean', was $2n = 2x = 24$ (Fig. 10, A). Accession 4C1 showed 11 metacentric or submetacentric chromosome pairs and one acrocentric chromosome pair (Fig. 10, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 41.4 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 10, C. The genome size (1C value) of accession 4C1 was 3.5044 pg (Table 3).

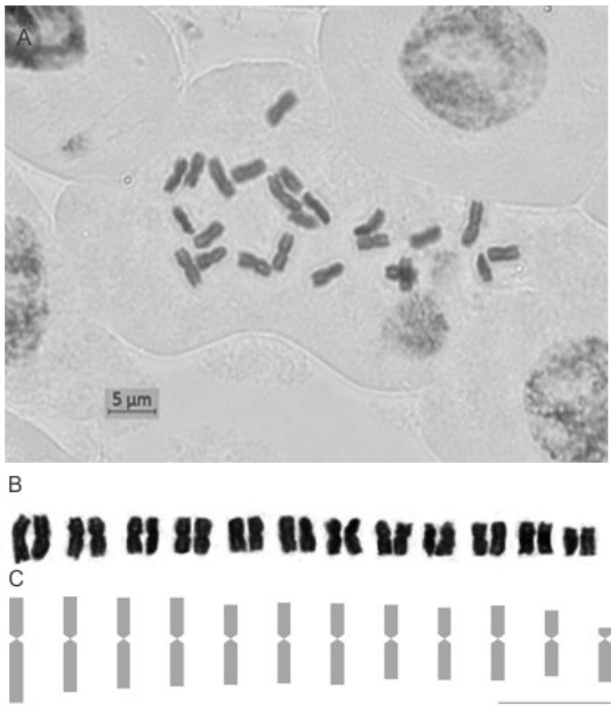

Fig. 11 – *C. chinense* variety 'Japonaise', accession 5C: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiograms; scalebar, 5 µm.

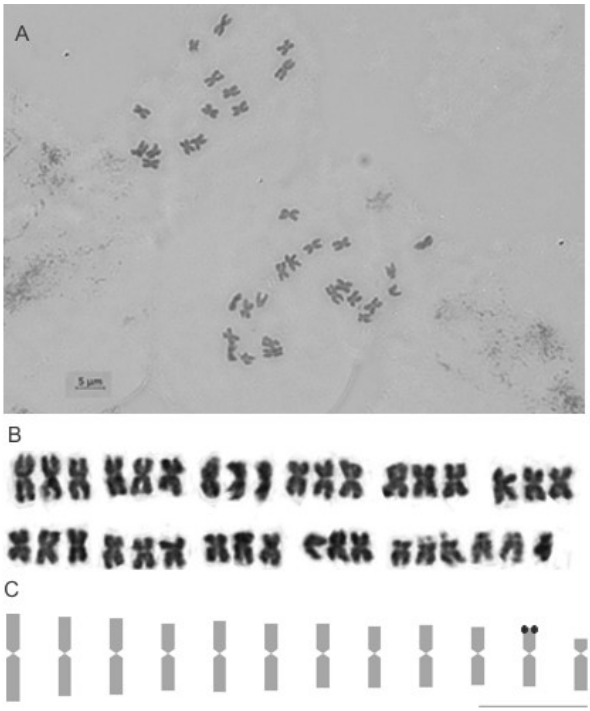

Fig. 12 – *C. chinense* variety 'Japonaise, accession 5C1: (A) metaphase plate with $2n = 3x = 36$ (B) karyotype (C) haploid idiogram with satellites (NORs) indicated as black full circles; scalebar, 5 µm.

The chromosome number for accession 5C, *C. chinense* 'Japonaise', was $2n = 2x = 24$ (Fig. 11, A). Accession 5C showed 11 metacentric or submetacentric chromosome pairs and one acrocentric chromosome pair (Fig. 11, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 42.7 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 11, C. The genome size (1C value) of accession 5C was 3.4367 pg (Table 3).

The chromosome number for accession 5C1, *C. chinense* 'Japonaise', was $2n = 3x = 36$ (Fig. 12, A). Accession 5C1 showed 10 metacentric or submetacentric chromosomes and two acrocentric chromosomes per haploid complement, with a satellite on chromosome 11 (Fig. 12, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 37.3 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 12, C. The genome size (1C value) of accession 5C1 was 5.2253 pg (Table 3).

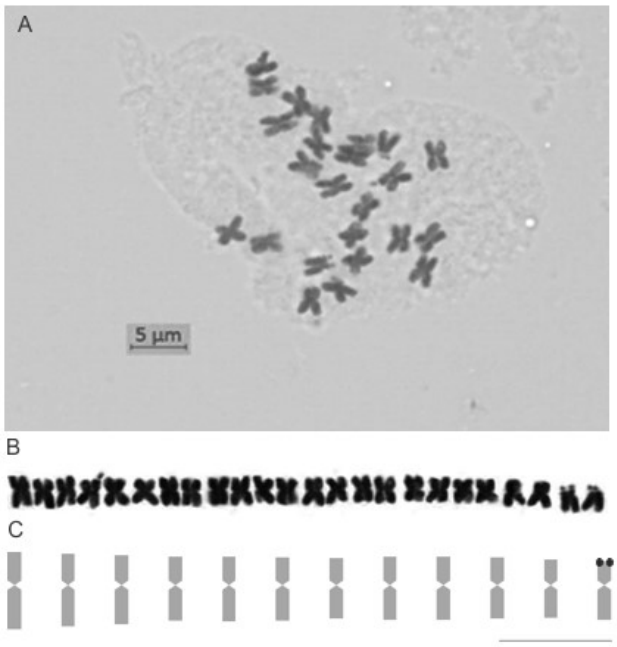

Fig. 13 – *C. chinense* variety 'Beni Highland', accession 6C: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram with satellites (NORs) indicated as black full circles; scalebar, 5 µm.

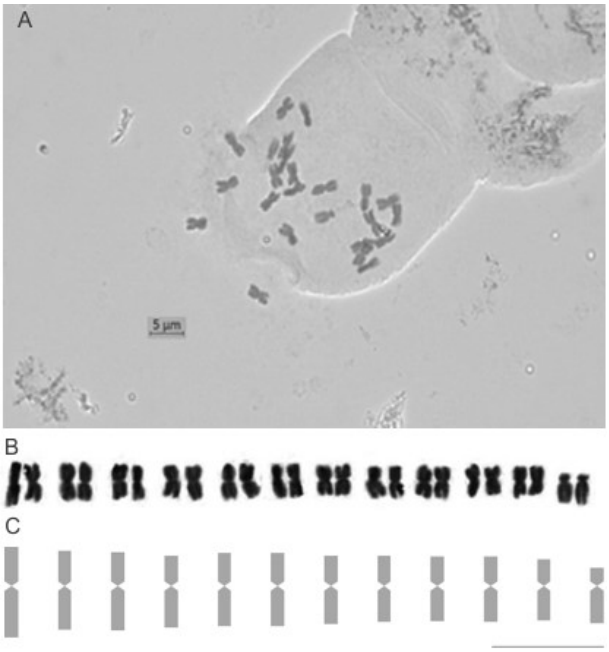

Fig. 14 – *C. chinense* variety 'Beni Highland', accession 6C1: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram; scalebar, 5 µm.

The chromosome number for accession 6C, *C. chinense* 'Beni Highland', was $2n = 2x = 24$ (Fig. 13, A). Accession 6C showed 11 metacentric or submetacentric chromosome pairs and one acrocentric chromosome pair carrying a satellite (Fig. 13, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 34.4 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 13, C. The genome size (1C value) of accession 6C was 3.4506 pg (Table 3).

The chromosome number for accession 6C1, *C. chinense* 'Beni Highland', was $2n = 2x = 24$ (Fig. 14, A). Accession 6C1 showed 11 metacentric or submetacentric chromosome pairs and one acrocentric chromosome pair (Fig. 14, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 37.6 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 14, C. The genome size (1C value) of accession 6C1 was 3.5151 pg (Table 3).

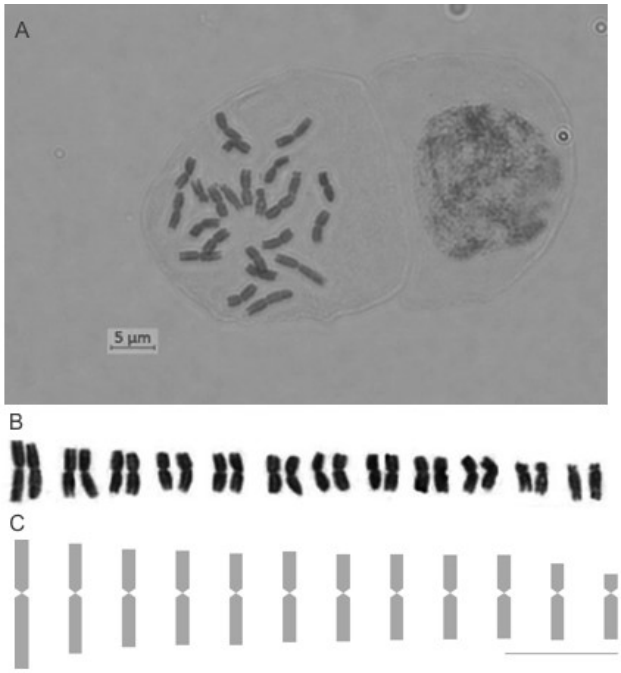

Fig. 15 – *C. chinense* variety 'Lippenstift', accession 7C: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram; scalebar, 5 µm.

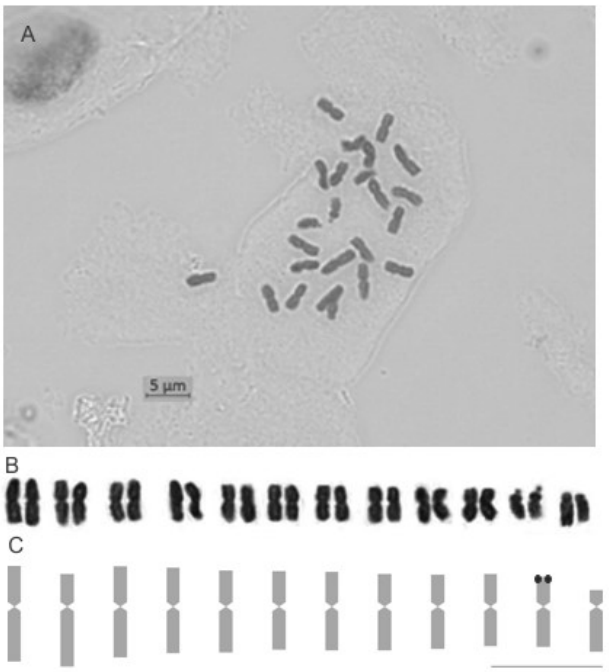

Fig. 16 – *C. chinense* variety 'Lippenstift', accession 7C1: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram with satellites (NORs) indicated as black full circles; scalebar, 5 µm.

The chromosome number for accession 7C, *C. chinense* 'Lippenstift', was $2n = 2x = 24$ (Fig. 15, A). Accession 7C showed 10 metacentric or submetacentric chromosome pairs, one submetacentric pair and one acrocentric chromosome pair (Fig. 15, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 48.8 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 15, C. The genome size (1C value) of accession 7C was 3.4707 pg (Table 3).

The chromosome number for accession 7C1, *C. chinense* 'Lippenstift', was $2n = 2x = 24$ (Fig. 16, A). Accession 7C1 showed 10 metacentric or submetacentric chromosome pairs, one submetacentric pair carrying a satellite and one acrocentric chromosome pair (Fig. 16, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 43.2 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 16, C. The genome size (1C value) of accession 7C1 was 3.4676 pg (Table 3).

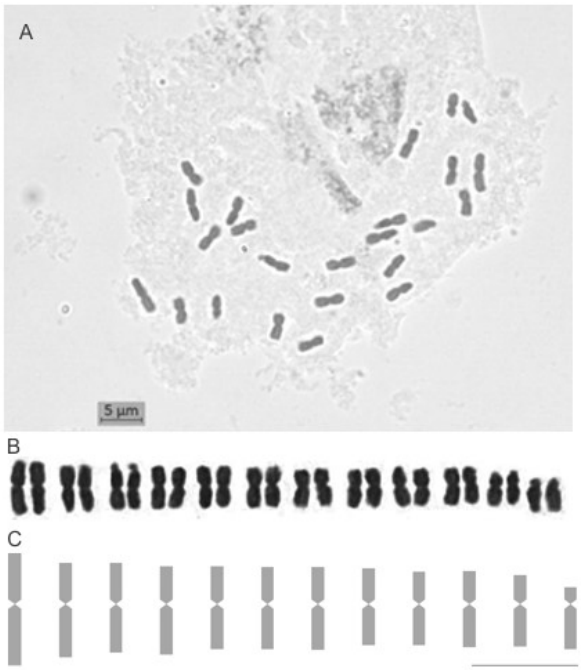

Fig. 17 – *C. chinense* variety 'Large Orange Thai', accession 8C: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram; scalebar, 5 µm.

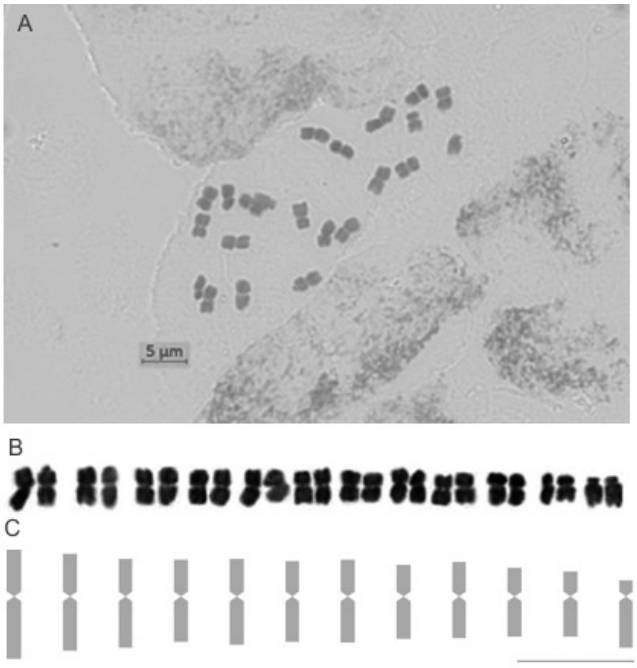

Fig. 18 – *C. chinense* variety 'Large Orange Thai', accession 8C1: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram; scalebar, 5 µm.

The chromosome number for accession 8C, *C. chinense* 'Large Orange Thai', was $2n = 2x = 24$ (Fig. 17, A). Accession 8C showed 11 metacentric or submetacentric chromosome pairs and one acrocentric chromosome pair (Fig. 17, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 46.6 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 17, C. The genome size (1C value) of accession 8C was 3.4536 pg (Table 3).

The chromosome number for accession 8C1, *C. chinense* 'Large Orange Thai', was $2n = 2x = 24$ (Fig. 18, A). Accession 8C1 showed 11 metacentric or submetacentric chromosome pairs and one acrocentric chromosome pair (Fig. 18, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 41.8 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 18, C. The genome size (1C value) of accession 8C1 was 3.4773 pg (Table 3).

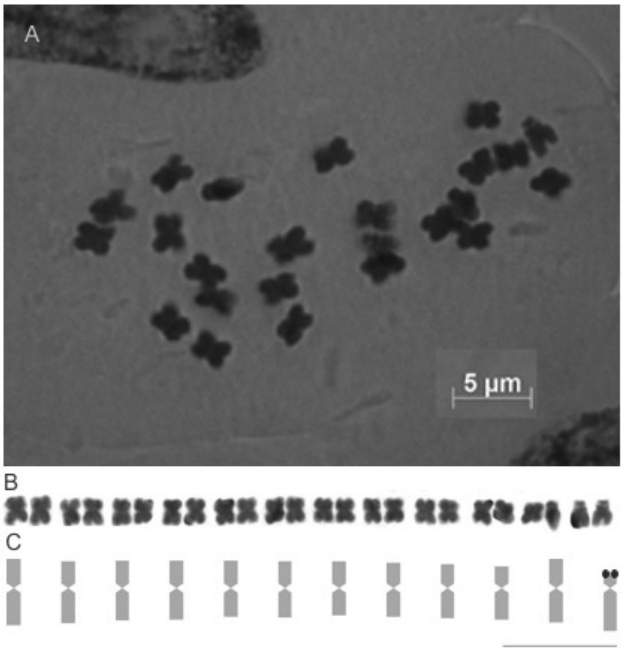

Fig. 19 – *C. annuum* variety 'Gänseschnabel', accession 9C: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram with satellites (NORs) indicated as black full circles; scalebar, 5 µm.

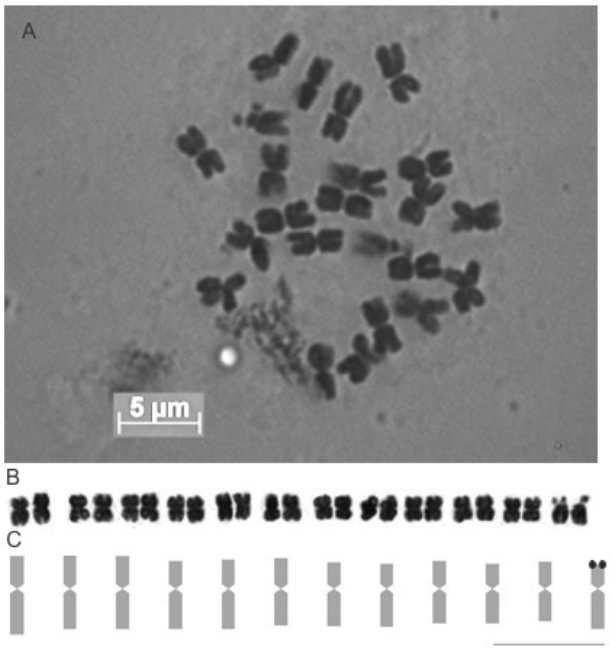

Fig. 20 – *C. annuum* variety 'Gänseschnabel', accession 9C1: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram with satellites (NORs) indicated as black full circles; scalebar, 5 µm.

The chromosome number for accession 9C, *C. annuum* 'Gänseschnabel', was $2n = 2x = 24$ (Fig. 19, A). Accession 9C showed 10 metacentric or submetacentric chromosome pairs and two acrocentric chromosome pairs, one of them possibly carrying a satellite (Fig. 19, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 31.0 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 19, C. The genome size (1C value) of accession 9C was 3.4283 pg (Table 3).

The chromosome number for accession 9C1, *C. annuum* 'Gänseschnabel', was $2n = 2x = 24$ (Fig. 20, A). Accession 9C1 showed 11 metacentric to submetacentric chromosome pairs and one acrocentric chromosome pair carrying a satellite (Fig. 20, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 36.5 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 20, C. The genome size (1C value) of accession 9C1 was 3.4081 pg (Table 3).

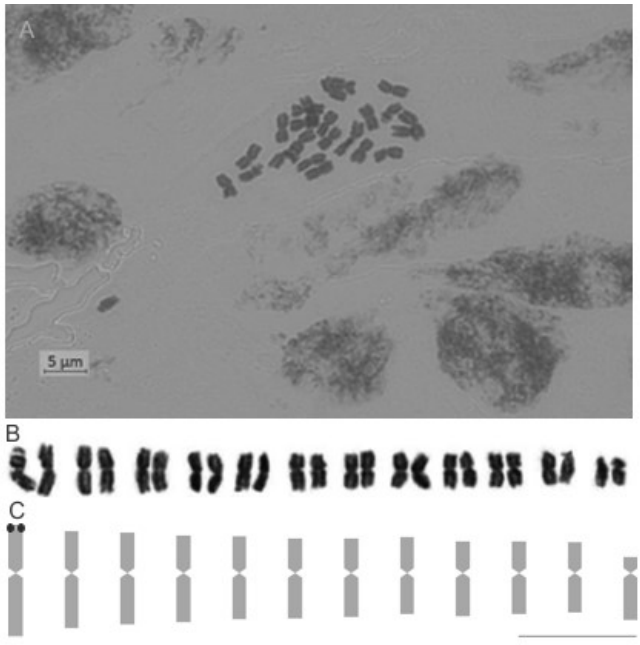

Fig. 21 – *C. annuum* variety 'Apfelpaprika', accession 10C: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram with satellites (NORs) indicated as black full circles; scalebar, 5 µm.

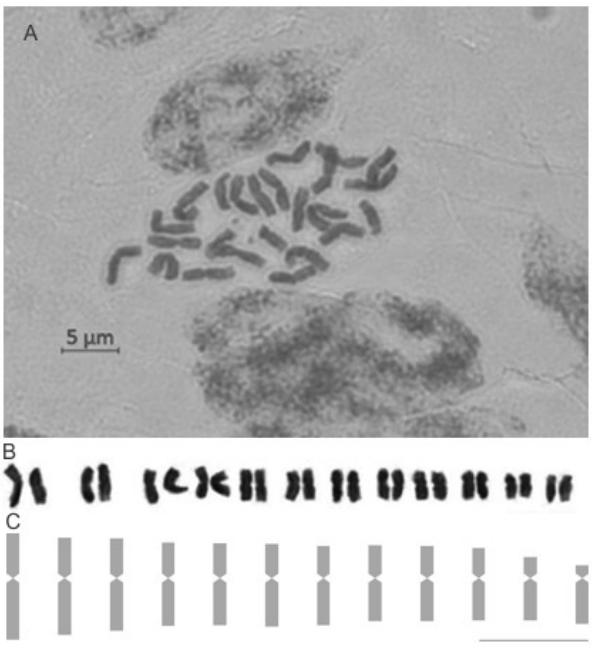

Fig. 22 – *C. annuum* variety 'Apfelpaprika', accession 10C2: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram; scalebar, 5 µm.

The chromosome number for accession 10C, *C. annuum* 'Apfelpaprika', was $2n = 2x = 24$ (Fig. 21, A). Accession 10C showed 11 metacentric or submetacentric chromosome pairs, the largest of them carrying a satellite, and one acrocentric chromosome pair (Fig. 21, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 42.2 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 21, C. The genome size (1C value) of accession 10C was 3.4162 pg (Table 3).

The chromosome number for accession 10C2, *C. annuum* 'Apfelpaprika', was $2n = 2x = 24$ (Fig. 22, A). Accession 10C2 showed 11 metacentric or submetacentric chromosome pairs and one acrocentric chromosome pair (Fig. 22, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 44.5 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 22, C. The genome size (1C value) of accession 10C2 was 3.4283 pg (Table 3).

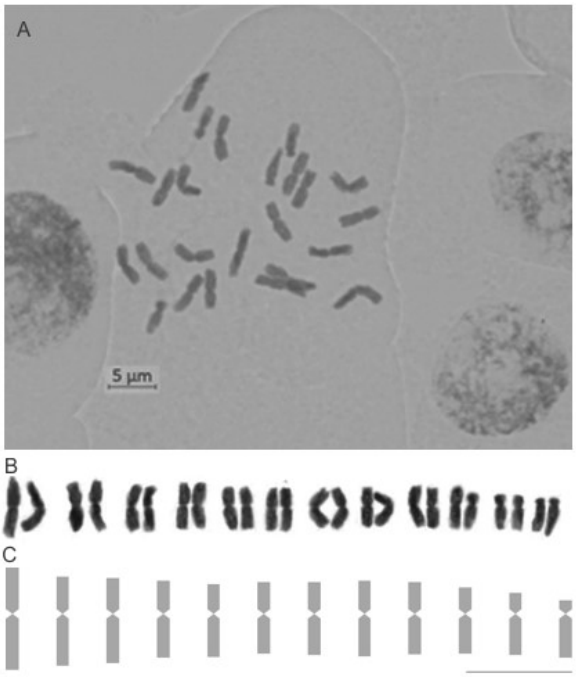

Fig. 23 – *C. annuum* variety 'Jalapeno', accession 11C: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram; scalebar, 5 µm.

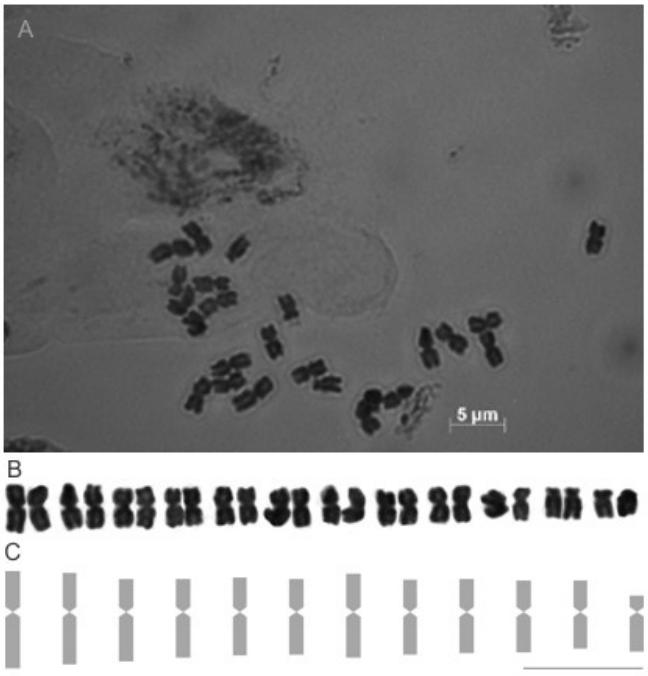

Fig. 24 – *C. annuum* variety 'Jalapeno', accession 11C1: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram; scalebar, 5 µm.

The chromosome number for accession 11C, *C. annuum* 'Jalapeno', was $2n = 2x = 24$ (Fig. 23, A). Accession 11C showed 10 metacentric or submetacentric chromosome pairs and two acrocentric chromosome pairs (Fig. 23, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 42.3 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 23, C. The genome size (1C value) of accession 11C was 3.4249 pg (Table 3).

The chromosome number for accession 11C1, *C. annuum* 'Jalapeno', was $2n = 2x = 24$ (Fig. 24, A). Accession 11C1 showed 10 metacentric or submetacentric chromosome pairs and two acrocentric chromosome pairs (Fig. 24, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 38.9 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 24, C. The genome size (1C value) of accession 11C1 was 3.4622 pg (Table 3).

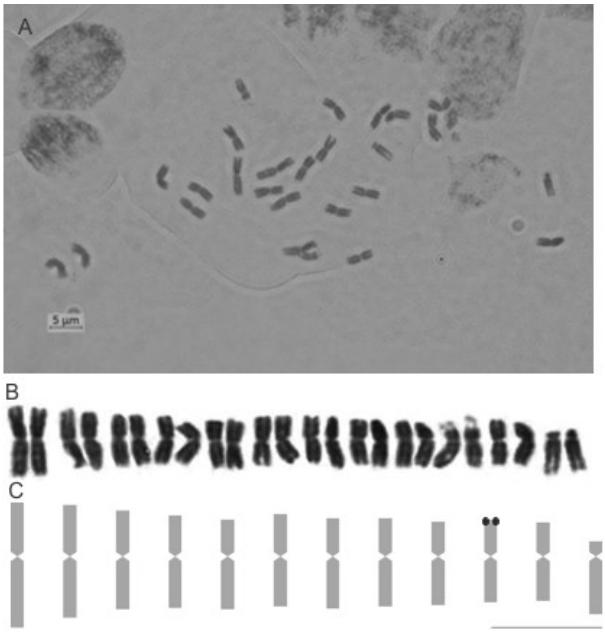

Fig. 25 – *C. chinense* variety 'Jalapeno', accession 11C2: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram with satellites (NORs) indicated as black full circles; scalebar, 5 µm.

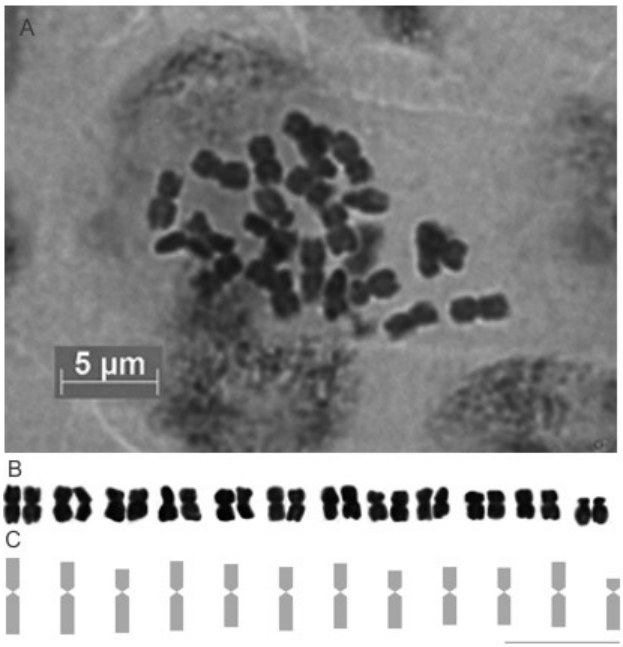

Fig. 26 – *C. chinense* variety 'Habanero', accession 12C1: (A) metaphase plate with $2n = 2x = 24$ (B) karyotype (C) haploid idiogram; scalebar, 5 µm.

The chromosome number for accession 11C2, *C. annuum* 'Jalapeno', was $2n = 2x = 24$ (Fig. 25, A). Accession 11C2 showed 10 metacentric or submetacentric chromosome pairs, one submetacentric pair carrying a satellite, and one acrocentric chromosome pair (Fig. 25, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 50.0 µm (Table x). The corresponding idiogram based on the measurements is given in Fig. 25, C. The genome size (1C value) of accession 11C2 was 3.4094 pg (Table 3).

The chromosome number for accession 12C1, *C. chinense* 'Habanero', was $2n = 2x = 24$ (Fig. 26, A). Accession 12C1 showed 11 metacentric or submetacentric chromosome pairs and one acrocentric chromosome pair (Fig. 26, B). Measurements of the chromosome lengths resulted in a total haploid chromosome length of 33.4 µm (Table 3). The corresponding idiogram based on the measurements is given in Fig. 26, C. The genome size (1C value) of accession 12C1 was 3.4449 pg (Table 3).

5.3. Fluorescent *in situ* hybridization

To characterize some accessions in more detail, fluorescent *in situ* hybridization with two rDNA probes (5S and 18S rDNA – part of 45S rDNA unit) was performed in three individuals, triploid *Capsicum chinense* 'Japonaise' (5C1), diploid *C. chinense* 'Billy Goat' (1C1) and diploid *C. annuum* 'Apfelpaprika' (10C). The 5S and 18S rRNA gene loci were localized on chromosomes of the two analyzed species to examine and compare loci number and their position. Two subterminal signals of 5S rDNA on one chromosome pair and four subterminal signals of 18S rDNA on two chromosome pairs were detected in diploid accession 1C1 of *C. chinense* (Figure 27, B). In the triploid individual of *Capsicum chinense* 'Japonaise' (5C1) three strong 5S rDNA signals were detected in subterminal regions on the short arms of three homologous metacentric chromosomes (Figure 28, red). Six hybridization signals of the 18S rDNA probe were observed, three of which appeared at the terminal region of the short arms of small, acrocentric chromosomes. The other three 18S rDNA loci were within apparent satellites of three homologous larger, metacentric chromosomes (Figure 28, green).

The position of the 5S rDNA on chromosomes of diploid *C. annuum* 'Apfelpaprika' (10C) was subterminal in the short arms of one metacentric chromosome pair (Figure 27, A, shown in red). The hybridization of the 18S rDNA probe resulted in two strong hybridization signals in subterminal positions of the short arms of a pair of smaller metacentric chromosomes, associated with sattelites (Figure 27, A, green).

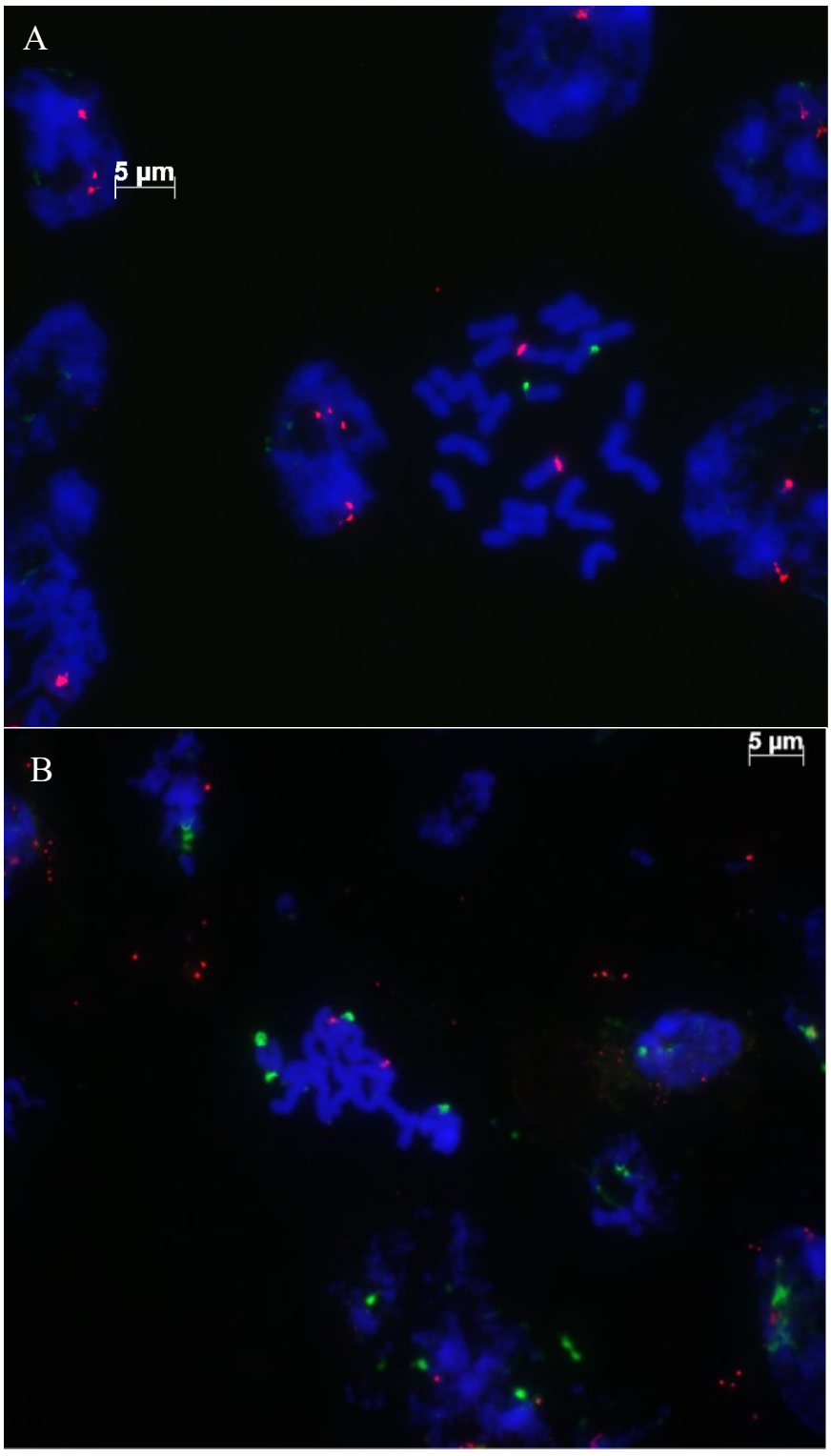

Fig. 27– FISH with 18S rDNA (green) and 5S rDNA (red) probes mapped on somatic chromosomes of (A) *C. annuum* cv. 'Apfelpaprika' (2*n* = 24; individual 10C); (B) *C. chinense* cv. 'Billy Goat' (2*n* = 24; individual 1C1)

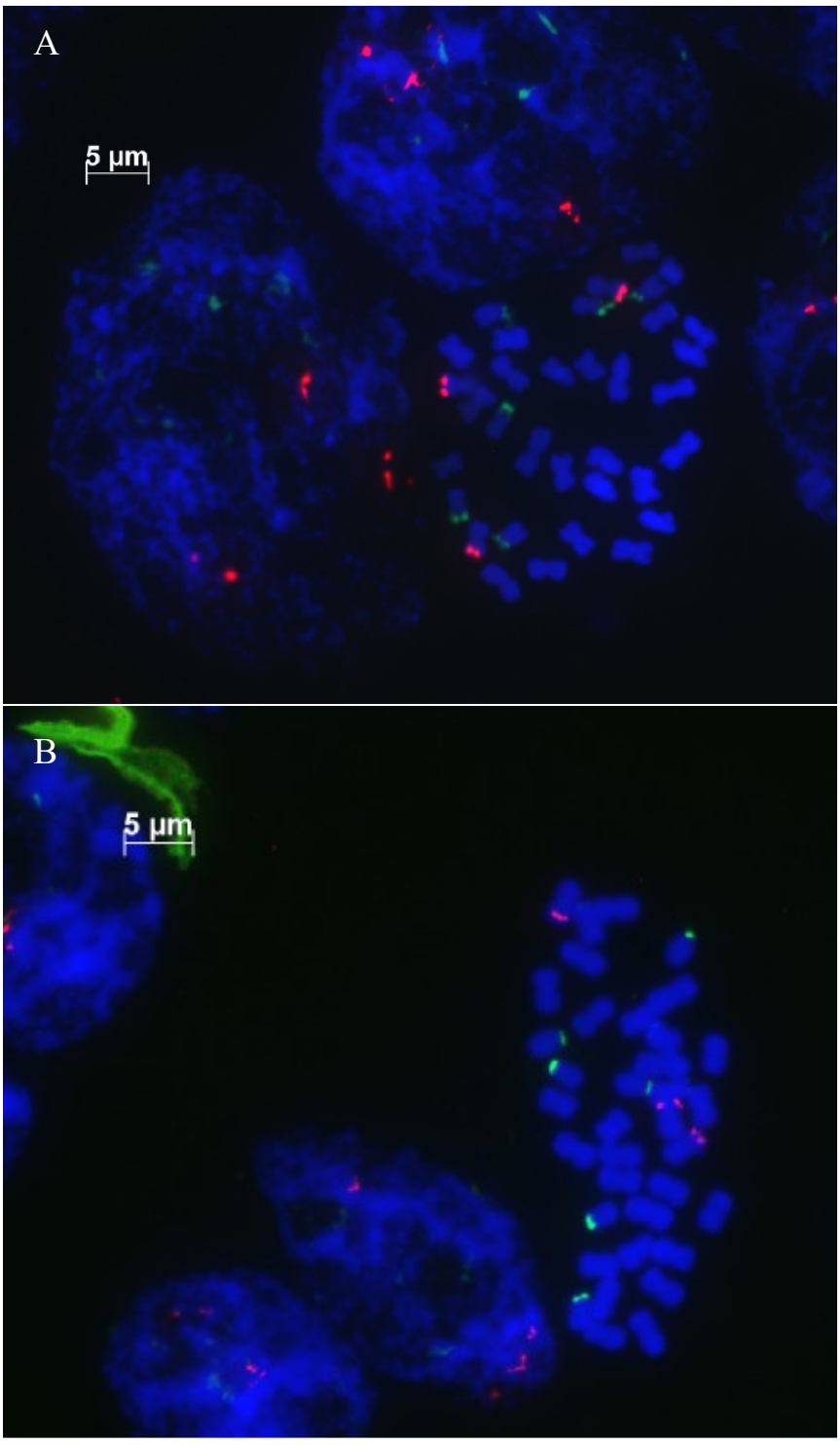

Fig. 28 – FISH with 18S rDNA (green) and 5S rDNA (red) probes mapped on somatic chromosomes of *C. chinense* cv. 'Japonaise' ($2n = 36$; individual 5C1; A, B).

5.4. Next-generation sequencing

Repetitive DNAs of *Capsicum annuum* and *C. chinense* were analyzed using the RepeatExplorer pipeline. Clustering of reads was performed individually for each species, followed by a comparative analysis using reads from both species. Table 4 provides a summary of the genomic proportions of the repetitive elements in the genomes of *C. annuum* and *C. chinense*. Additionally, Figure 29 shows a bar graph including the proportions of the identified repeat types for each species.

In both analyzed species, repeats identified as LTR (Long terminal repeat)-retrotransposons constituted the largest fraction of the total genomic composition. In *C. annuum* and *C. chinense*, significantly higher amounts of *Ty3-gypsy* elements were found relative to *Ty1-copia* elements. The majority of repeats originated from the Chromovirus lineage in *C. annuum* (18.2 %), while the other two *Ty3-gypsy* lineages (Athila, Ogre-Tat) each represented about 3% of the genome. In *C. chinense*, Athila repeats were significantly more frequent (16.6%), while the other lineages of the *Ty3-gypsy* superfamily were at similar proportions as in *C. annum* (Table 4). Of the *Ty1-copia* superfamily, which comprised around 5% of the genome in both species, TAR elements were most frequent, with proportions of 1.6% in *C. annuum* and 1.9% in *C. chinense*.

No Helitrons were detected across genomes of both analyzed *Capsicum* species. Moderate amounts of DNA transposons were present in *C. annuum,* the lineages combined representing only 1.7% of the genome. In *C. chinense,* those lineages were more frequently represented, reaching a genome proportion of 3.4%. Of these lineages, most DNA transposons were identified as Mutator type in both analyzed species (Table 4).

Of the non-LTR retrotransposons, LINEs (Long interspersed nuclear elements) made up 1.7% in *C. annuum* and 1.4% in *C. chinense*. Other types of dispersed repeats, such as Pararetroviruses reached proportions of 3.4% and 3.7% of the genomes in *C. annuum* and *C. chinense* respectively. MITEs (Miniature inverted-repeat transposable elements) were not identified in either species.

Of the tandem repeats, rDNAs made up 3% of the genome in *C. annuum* and 3.8% in *C. chinense*. Satellite DNAs reached a genome proportion of 2.1% in *C. annuum* and 0.9% in *C. chinense*. 58% and 42.7% of the genomes remained unclassified in *C. annuum* and *C. chinense*, respectively.

Four clusters were identified as putative satellites in the genomes of *C. annuum* and *C. chinense* each. The structures of the satellite repeat motifs were further analyzed using DOTTER (Sonnhammer & Durbin 1995). Based on the shape of the cluster graph (Novák et al. 2010) and the presence of tandem repeats in head-to-tail orientation in assembled contigs (Fig. 30), the clusters were classified as satellite DNAs. The number of reads as well as the genomic proportion of each cluster is provided in Table 5. *k*-mer frequency analysis (Macas et al. 2010) was used with 25 bp long *k*-mers for consensus sequence reconstruction, which resulted in monomers of 37 bp to 112 bp in length (Table 5). Cluster CL112 is a higher order satellite repeat and thus is represented by consensus submonomer (length: 33 bp). Reconstructed consensus sequences are provided as sequence logos (Schneider & Stephens 1990; Figs. 31-38), and summarized based on relative frequencies in Table 5. Graphs of the satellite DNAs are provided in Figures 31-38.

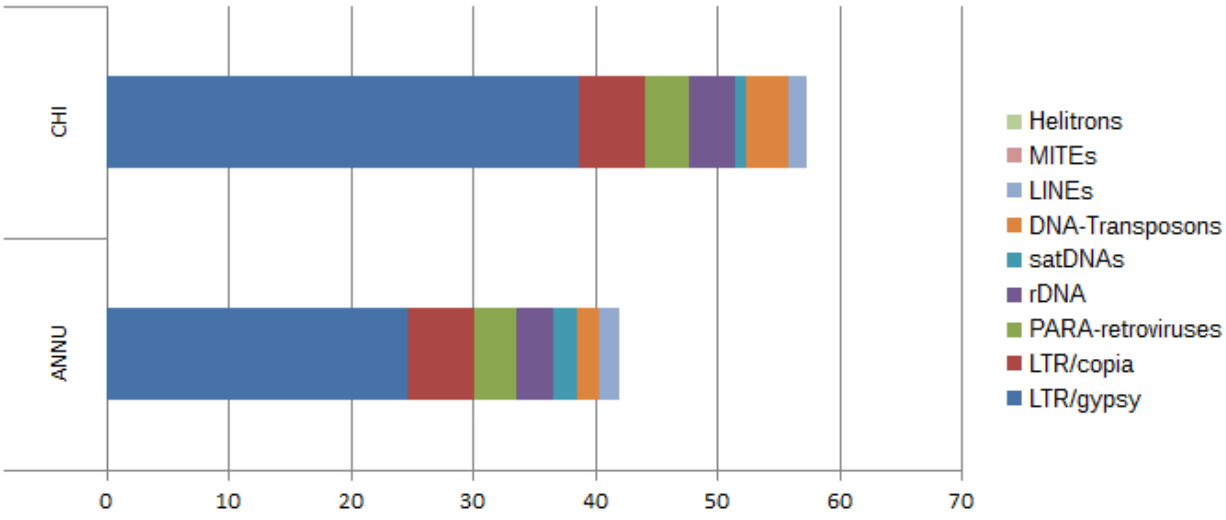

Figure 29 – Proportions (in %) of repetitive elements for each species, colored by repeat type (see legend). The species are abbreviated as follows: ANNU = *Capsicum annuum*; CHI = *C. chinense*;

Two satellites of *C. annuum,* referred to as CL103 and CL105, showed partial homology (Fig. 30). The satellites referred to as CL81 in *C. annuum* and CL 44 in *C. chinense* appear to be homologous as well, based on the dot plot analysis (DOTTER, Fig. 30) and the identical monomer length (37bp). Additionally, CL120 in *C. annuum* is homologous to CL101 in *C. chinense* (Fig. 30), both monomers showing a consensus length of 112bp.

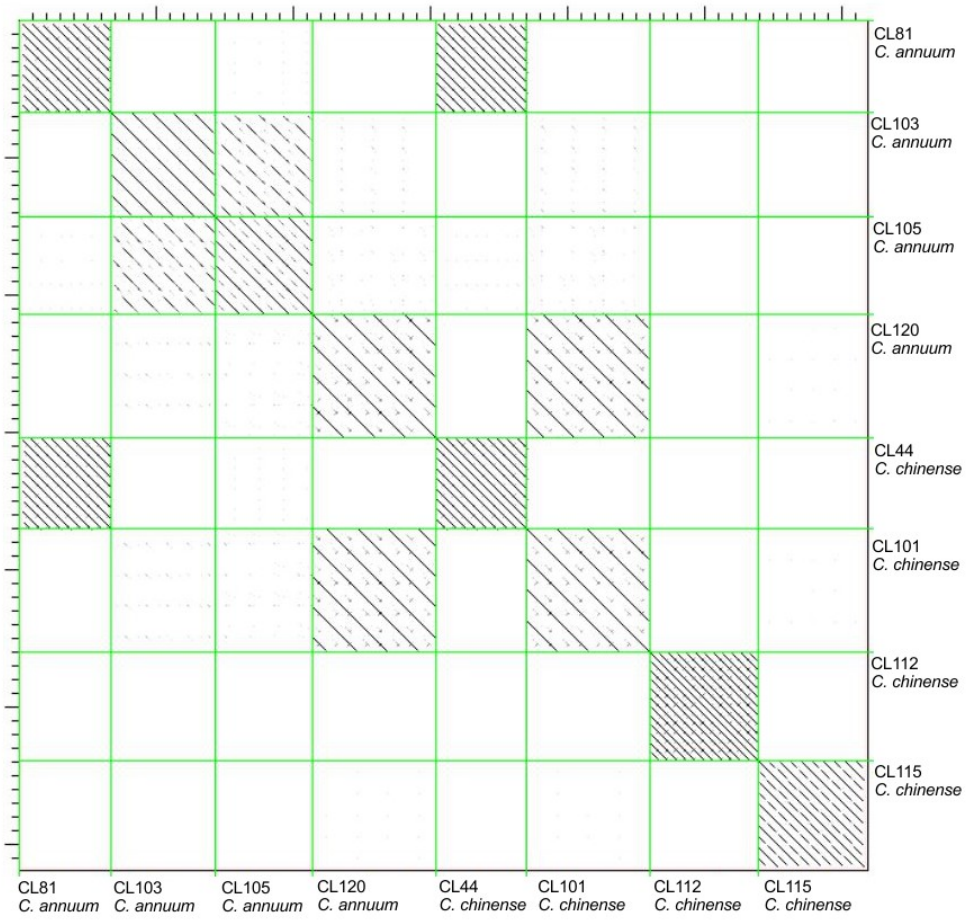

Figure 30 – Comparison of the elements identified as satellite DNA of *Capsicum annuum* and *C. chinense* by dot-plot analysis. The vertical and horizontal axes represent the sequences of the eight monomers divided by green lines.

Table 4 – Estimates of the genome proportion (%) of various repeat types identified in the analyzed diploid genomes of *Capsicum annuum* and *C. chinense*

Type	Repeat family	*C. annuum*		*C. chinense*	
		reads	%	reads	%
copia	-	188381	5.0	170256	5.4
	Ale	14699	0.4	18449	0.6
	Alesia	0	0.0	0	0.0
	Angela	2119	0.1	835	0.0
	Bianca	35706	1.0	30996	1.0
	Ikeros	9817	0.3	8685	0.3
	Ivana	9425	0.3	6945	0.2
	Sire	43638	1.3	22872	0.7
	Tar	56484	1.6	58610	1.9
	Tork	16493	0.5	22864	0.7
gypsy	-	847190	24.7	1211816	38.7
	Athila	112980	3.3	521320	16.6
	Chromo	625107	18.2	589354	18.8
	Ogre_Tat	109103	3.2	101142	3.2
Helitron	-	1134	0.0	0	0.0
DNA transposons	-	59328	1.7	107420	3.4
	Cacta	11145	0.3	8282	0.3
	Harbinger	0	0.0	3088	0.1
	Hat	17323	0.5	29960	1.0
	Mariner	0	0.0	10129	0.3
	Mutator	30860	0.9	55961	1.8
Other/ nonLTR	-	173500	5.1	160772	5.1
	LINE	57691	1.7	45352	1.4
	MITE	0	0.0	0	0.0
	PARA	115809	3.4	115420	3.7
Tandem repeats	-	173686	5.1	39322	3.8
	rDNA	102885	3.0	118162	3.8
	satDNA	70801	2.1	27460	0.9
Unclassified	-	1989639	58.0	1338828	42.7

Table 5 – Characterization of satellite DNA repeats in diploid species *C. annuum* and *C. chinense*

Species	Cluster	Genome proportion (%)	Number of reads	Consensus length (bp)	Consensus structure
C. annuum	CL81	0.21	13369	37	TGGGATTTTACTTTCTGTTCAGG ACCCTCCTGAAAAA
	CL103	0.13	8075	76	TAAAGGGGATTTTTTTGTTTTGA ATGTGGTTTGGTTGGTGACTTA GTTACTTGAAAATGGTTAAGTG TAGGGACTT
	CL105	0.12	7745	44	CCATTTTCAAGTCACTAATTCAC ATTTCAAGTAACTACACTTAA
	CL120	0.096	6069	112	CTTGAGCTAAGTCTTGCCTCTA AGGCAAGAGTTGTAGAACATCT CATAAATGAAAACATAATGTGG TAGTGTCAACTCTTGCCCGTGG GGCATAATTTGTTGTTTGAAAG TC
C. chinense	CL44	0.33	18583	37	TGGGATTTTACTTTCTGTTCAGG ACCCTCCTGAAAAA
	CL101	0.087	4984	112	TGAAAACATAATGTGGTAGTGT CAACTCTTGCCCGTGGGGCATA ATTTGTTGTTTGAAAGTCCTTGA GCTAAGTCTTGCCTCTAAGGCA AGAGTTGTAGAACATCTCATAA A
	CL112	0.069	3940	33	GCGGCGCCTAACCTATCGCCTA TGTTGACATAG
	CL115	0.066	3760	79	ACTAGGGGATGACGCCCTATGA AGACGAATAAGATGTCAACTAG GGGATGGCGCCCTATGAAGACG AGATAAGATGTCA

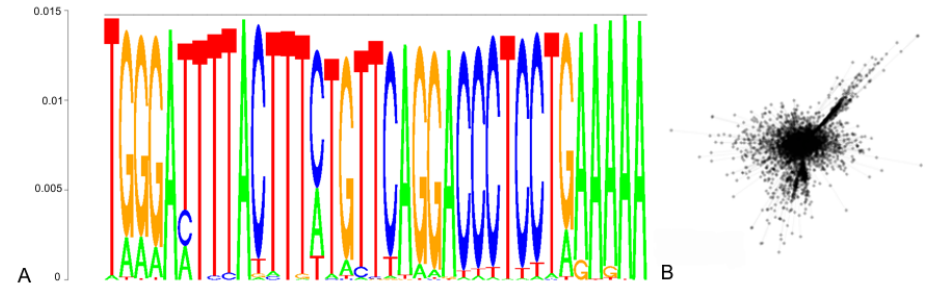

Figure 31 – Monomer characterization of cluster CL81, *C. annuum*. (A) Monomer sequence logo; (B) Graph layout identified by NGS.

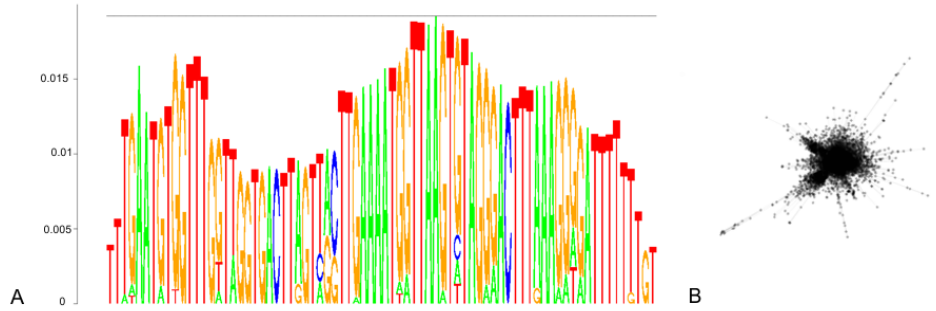

Figure 32 – Monomer characterization of cluster CL103, *C. annuum*. (A) Monomer sequence logo; (B) Graph layout identified by NGS.

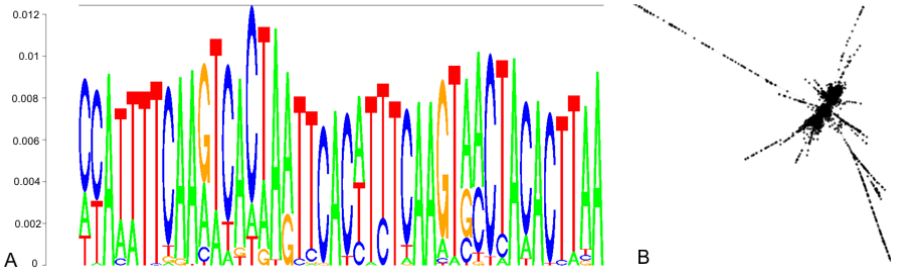

Figure 33 – Monomer characterization of cluster CL105, *C. annuum*. (A) Monomer sequence logo; (B) Graph layout identified by NGS.

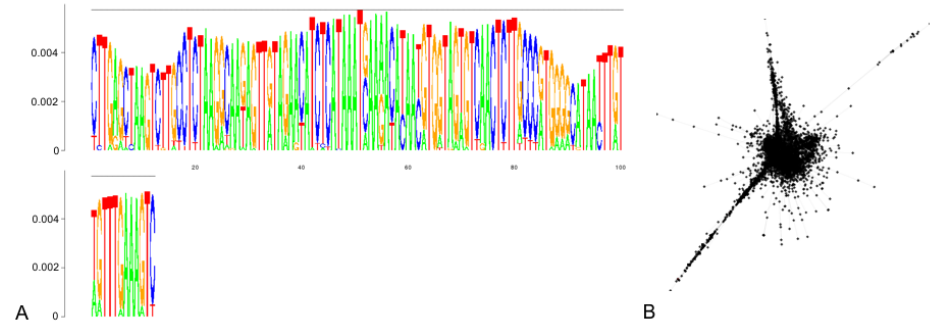

Figure 34 – Monomer characterization of cluster CL120, *C. annuum*. (A) Monomer sequence logo; (B) Graph layout identified by NGS.

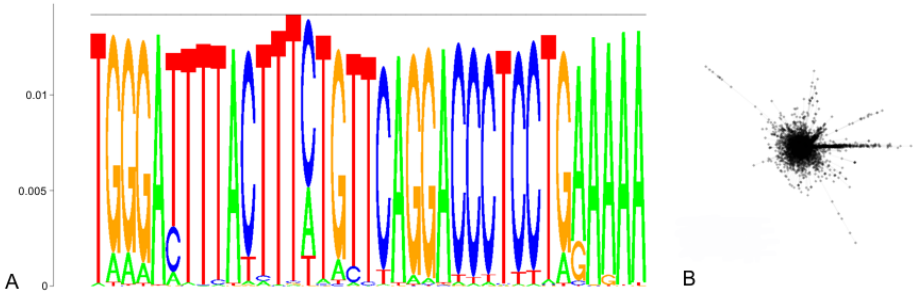

Figure 35 – Monomer characterization of cluster CL44, *C. chinense*. (A) Monomer sequence logo; (B) Graph layout identified by NGS.

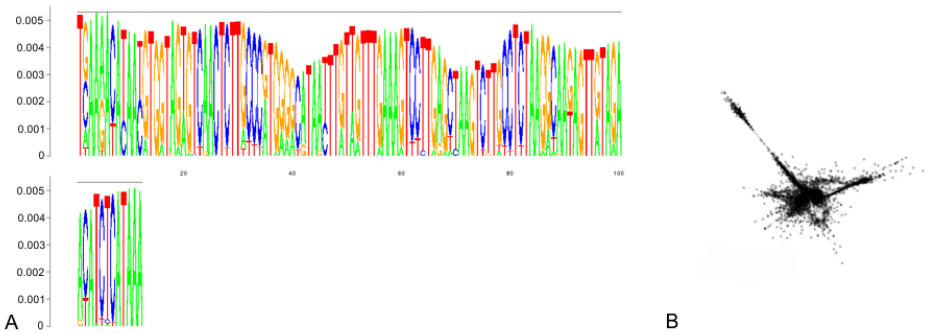

Figure 36 – Monomer characterization of cluster CL101, *C. chinense*. (A) Monomer sequence logo; (B) Graph layout identified by NGS.

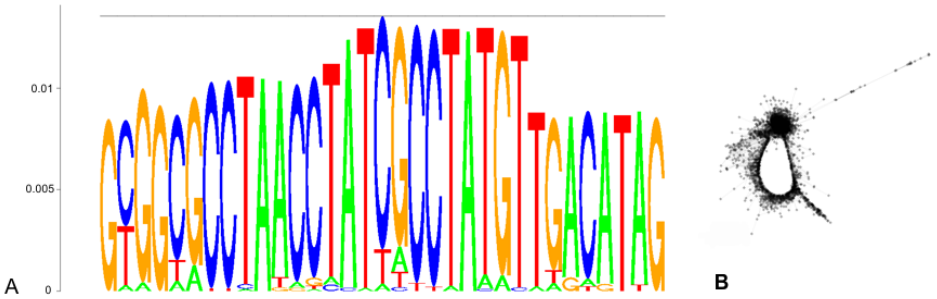

Figure 37 – Partial submonomer characterization of cluster CL112, *C. chinense*. (A) Monomer sequence logo; (B) Graph layout identified by NGS.

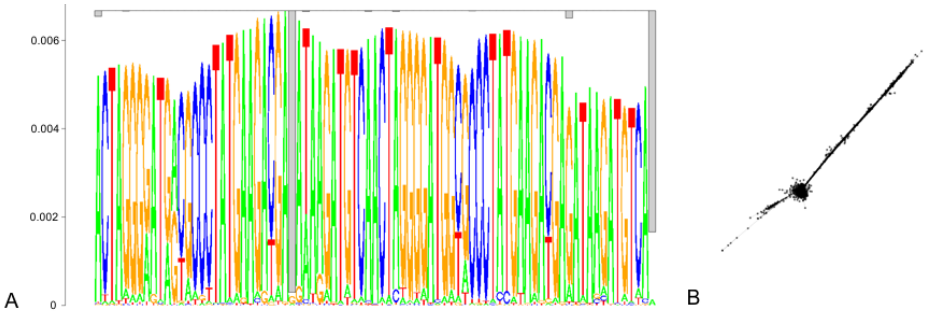

Figure 38 – Monomer characterization of cluster CL115, *C. chinense*. (A) Monomer sequence logo; (B) Graph layout identified by NGS.

6. DISCUSSION

6.1. Flow cytometry

The measurements using flow cytometry revealed relatively similar genome size values throughout the analyzed varieties and individuals, ranging from 1C = 3.41 pg to 3.58 pg for the diploid accessions. Slight fluctuations could be caused by changes in technical settings such as the speed of the flow cytometer, which influences the duration of the measurements. Only the DNA content of accession 5C1 (*C. chinense* 'Japonaise') was significantly higher at 1C = 5.23 pg. As this number was about one third higher than the DNA content of the diploid accessions, it was suspected that 5C1 was possibly a triploid, which had to be confirmed by examining the mitotic metaphase plates and assessing the chromosome number. Comparing the present results with previous studies had to be done with caution, as data from other literature corresponded to different samples and cultivars. Moscone et al. (2003) published flow cytometric measurements for several *Capsicum* species and repeated similar genome size values, despite differences in methodology, variables during measurement, dye used for nuclei staining (ethidium bromide) and standard used (*Hordeum vulgare* L. 'Ditta'). According to this study, the mean 1C values for *C. annuum* accessions was 3.83 pg, and for *C. chinense* 4.02 pg. The present data therefore supports the claim that there is little intraspecific variation of DNA amount in *Capsicum* (Moscone et al. 2003), reflected also in the constant chromosome number and young age of these species. Thus, the difference in DNA amount between *C. annuum* and *C. chinense* accessions was found to be minimal.

6.2. Chromosome counts

Studying the chromosome numbers and morphology of *Capsicum* accessions provides important data for classification and taxonomy of cultivars. It facilitates the understanding of the developmental process from wild progenitors to cultivated plants and thus, by providing insights into genetic variation, helps uncover the evolution of the genus (Sousa et al. 2015). Chromosomes of the accessions were analyzed after Feulgen staining. Counting presented few difficulties because of the relatively large chromosome size of the accessions, the total chromosome length ranging from 2.2 to 5.8 µm. In comparison with other angiosperm chromosome sizes, which vary from 0.6 to over 14.6 µm (Weiss-Schneeweiss & Schneeweiss 2013), *Capsicum* chromosomes are considered medium-sized.

All species counted so far in the genus *Capsicum* are exclusively diploid – which is relatively common in Solanaceae in general (Pickersgill 2007). The reported chromosome numbers for the two species analyzed here were $2n = 2x = 24$ (Moscone et al. 1995; Pickersgill 1997; Scaldaferro et al. 2006). These results were also confirmed here for 23 of the 24 analyzed individuals. Although the chromosome number of $x = 13$ has also been recorded for the genus, it has only been found in a small number of wild species and was therefore not expected in this analysis (Moscone et al. 2007; Scaldaferro et al. 2006; Smith & Heiser 1951). The extra chromosome (13^{th}) is hypothesized to originate from centric (Robertsonian) fission of a large metacentric chromosome (Moscone et al. 2007). Karyotype symmetry is a common feature in angiosperms, which are known to mainly possess meta- and submetacentric chromosomes of similar lengths (Weiss-Schneeweiss & Schneeweiss 2013). Most *Capsicum* species (especially those with $2n = 24$) exhibit relatively similar karyotypes of 11 metacentric pairs and one subtelocentric pair of decreasing size, despite genome size varying up to 3-fold (Moscone et al. 2003). One of the metacentric chromosomes usually had a secondary constriction with a nucleolus organizing region (NOR), and the subtelocentric chromosomes frequently carried satellites (Kwon & Kim 2009). Both of these structures contain 45S rDNA genes. In *C. annuum* in particular, 10 metacentric to submetacentric chromosomes and two acrocentric chromosomes are common (Park et al. 1999). This is also found in the current study: the *C. annuum* accessions analyzed within this thesis also showed 10 metacentric to submetacentric chromosomes of similar, slightly decreasing size. The *C. annuum* accessions examined generally had two significantly smaller, acrocentric chromosome pairs, one of them carrying a satellite. More satellites on larger chromosomes could be inferred from the karyograms as well. Accession 9C1 (*C. annuum* 'Gänseschnabel') showed 11 metacentric to submetacentric chromosome pairs and one acrocentric chromosome pair, but this discrepancy relative to the other *annuum* accessions was likely due to the quality of chromosome spreads and their condensation (Fig. 20, A). The same was true for accessions 10C and 10C2 (*C. annuum* 'Apfelpaprika'), which showed 11 metacentric to submetacentric chromosome pairs and one acrocentric chromosome pair. Similar results were obtained for *C. chinense* cultivars, except for accession 5C1. The diploid *C. chinense* accessions showed 11 metacentric to submetacentric chromosome pairs gradually decreasing in size, one or two of them carrying satellites. Only one smaller acrocentric chromosome pair with a satellite was present.

High karyotype constancy was reported for the cultivated *Capsicum* species, as well as minor differences between *C. annuum* and *C. chinense* involving number of NORs, size of NOR-related satellites and heterochromatin amounts (Moscone et al. 2007). Both species belong to the Annuum clade, a complex with high genetic uniformity (Carrizo García et al. 2016). Based on the homogeneity of karyotypes across the $2n = 24$ *Capsicum* species, it was suggested that few chromosome rearrangements and genome reorganization happened during the genus' evolution (Moscone et al. 2003), especially during the domestication of the species, which often involved their hybridization, and thus, introgression. Therefore, karyotype symmetry was retained by potentially uniform DNA addition over both chromosomal arms (Moscone et al. 2003). Cheema and Pant (2013), who analyzed seven varieties of *C. annuum,* also confirm karyotype similarity (chromosome size, karyotype symmetry and chromosome types) across the varieties, but report differences in total haploid chromosome length.

Figure 39 – Flowers and fruits of diploid and triploid individuals of *C. chinense* var. 'Japonaise'. (A, C) flowers and fruits of diploid 5C individual; (B, D) flowers and fruits of triploid accession 5C1.

One of the 24 individuals that were examined as part of the present thesis did not possess the expected chromosome number of $2n = 2x = 24$. Individual 5C1 of *C. chinense* 'Japonaise' had consistently in all cells 36 instead of 24 chromosomes. Arranging the chromosomes into karyograms revealed triplets of homologous chromosomes, suggesting that the individual is an autotriploid, possibly due to a union of a reduced and unreduced gamete during the cultivation. Fluorescent *in situ* hybridization was performed to further characterize this individual. The chromosome number of 5C1 was therefore found to be $2n = 3x = 36$. Spontaneous triploids in *Capsicum* are an uncommon occurrence, and only a few instances of spontaneous and natural polyploidization in cultivated individuals have been reported by other authors (e.g. Chennaveeraiah & Habib 1972; Jha et al. 2012; Jha et al. 2017), mostly with polyploidy induced artificially (e.g. Kulkarni & Borse 2010; Zhang et al. 2003). So far, no polyploid individual has been found in wild *Capsicum* species (Carrizo García et al. 2016).

C. chinense accession 5C1 showed 10 metacentric to submetacentric chromosomes of similar size as well as two acrocentric chromosomes per haploid complement. One of the smaller subtelocentric chromosomes carried a satellite, as did possibly one or two of the larger chromosomes. The morphology of the triploid was largely the same as the diploid individual of the same cultivar (5C). Leaves, flowers and fruits of triploid accession 5C1 showed minor deviations from those of accession 5C (Fig. 39).

6.3. Fluorescent *in situ* hybridization

Individual chromosome pairs of *Capsicum* are similar in size and morphology and thus difficult to distinguish, especially between closely related species. Owing to this uniformity, conventional staining methods are not always helpful to resolve open questions (Romero-da Cruz et al. 2017). Fluorescent *in situ* hybridization is a molecular cytogenetic method which facilitates comparative studies of karyotypes to address evolutionary or taxonomic issues, as it allows more precise chromosome identification.

In the present thesis, FISH was used to further identify chromosome markers of three accessions, one of which was a spontaneous triploid of *C. chinense* (5C1). The present results are in accordance with findings from other studies (Kwon and Kim 2009), which obtained similar data for other *C.*

annuum accessions. Aguilera et al. (2016) detected a single 5S rDNA locus per haploid *Capsicum* genome. In the *C. chinense* triploid, there were two 45S rDNA loci (part of 45S rDNA monomer) in the short arms of all acrocentric chromosomes, i.e. in two acrocentric homologous chromosome triplets. In the diploid individual of *C. chinense* (1C1), 45S rDNA loci were found in two chromosome pairs. In *C. annuum*, only one of the acrocentric chromosome pairs carried the 45S rDNA locus. With the 45S rDNA located on the smallest acrocentric chromosomes as well as the 5S rDNA on larger, metacentric chromosomes in all three cases, the results for *C. annuum* and *C. chinense* are comparable and show a close relation between the two species.

45S rDNA loci frequently occur at subterminal chromosome positions because DNA rearrangements can happen without affecting gene order or linkages on the rest of the chromosome (Hansen et al. 1996, cited in Kwon & Kim 2009). The 5S rDNA locus on the interstitial position within a metacentric chromosome is considered to be a conserved feature of all *Capsicum* species (Scaldaferro et al. 2015). The present data confirm findings from previous analyses (Park et al. 1999; Park et al. 2000; Scaldaferro et al. 2006; Kwon & Kim 2009; Romero-da Cruz et al. 2016). Further, 45S rDNA loci (=18-5.8-26S rDNA) have been found to be highly variable (in number, size and position across different species of chile peppers, which may be the outcome of chromosomal rearrangements such as inversions, deletions or translocations, as well as from dispersal of rDNA repeats or amplification of satellite DNA originating from non-coding parts of the 45S rDNA region (Romero-da Cruz et al. 2017). In the two accessions analyzed, the 45S rDNA signals were detected on morphologically similar homoeologous chromosomes, which again asserts their phylogenetic closeness.

Lastly, because Feulgen staining and flow cytometry indicated accession 5C1 to be a triploid, localization of 5S and 45S rDNA probes allowed for further characterization of this individual. As predicted from examining the chromosome morphology, which showed triplets of similar chromosomes, the *in situ* hybridization confirmed the presence of three homologous chromosomes instead of chromosome pairs, as one additional 5S and 18S rDNA signal was present in the individual, as expected. Although further analysis – which would go beyond the scope of this thesis – is needed to answer questions concerning the polyploidization event, it can be hypothesized that the individual is likely of autopolyploid origin.

6.4. Next-generation sequencing

Repetitive DNAs, including tandem repeats and dispersed repeats (Weiss-Schneeweiss et al. 2015), account for a substantial fraction of plant genomes. *Capsicum* species generally show significant genome size diversity across the genus (Moscone et al.2003) but, with few exceptions (Moscone et al. 1995), share the same basic chromosome number of $x = 12$. It is therefore likely that the difference in genome sizes stems from variation in repetitive DNA amount (Park et al. 2012). Thus, chile pepper genomes are excellent models for the analysis of repetitive DNA evolution during speciation.

Next-generation sequencing was used here to characterize the profiles of repetitive DNA elements in genomes of two species of *Capsicum*, *C. annuum* and *C. chinense*. The two species analyzed here are phylogenetically closely related (Carrizo García et al. 2016). Both *C. annuum* and *C. chinense* are diploid, have similar nuclear haploid genome sizes (1C value; cf. Table 3) and the same chromosome number $2n = 2x = 24$.

In general, the repetitive DNA proportions in the two genomes were highly similar, probably owing to the close relation of *C. annuum* and *C. chinense*. In both analyzed species, transposable elements made up the largest fraction of repetitive DNA, which is a common occurrence in flowering plants (Tenaillon et al. 2010; Weiss-Schneeweiss et al. 2015). In the genomes of higher plants, *Ty3*-gypsy and *Ty1*-copia, two major families of LTR retrotransposons, are typically the dominant types of dispersed repeats (Kejnovský et al. 2012). The present analysis found *Ty3*-gypsy elements most abundant in both *C. annuum* (24.7%) and *C. chinense* (38.7%) genomes. *Ty1*-copia elements made up a smaller proportion of the genomes, with 5% and 5.4% in *C. annuum* and *C. chinense* respectively, but appeared to be more diverse with considerably larger number of different families. In both species, Chromoviruses constituted the major branch of *Ty3*-gypsy lineages with 18.2% in *C. annuum* and 18.8% in *C. chinense*. A previous study of constitutive heterochromatic regions in *C. annuum* found that *Ty3*-gypsy elements of the Del subgroup, which belongs to the chromoviruses, made up the largest part of repeats in heterochromatic regions, based on the sequence analyses of seven heterochromatin rich BAC clones (Park et al. 2012), which is in agreement with the current results.

Tandem repeats can be identified and efficiently characterized using NGS technology. Tandem repeats made up 5.1% of the genome of *C. annuum*, 2.1% of which were identified as satellite DNA. In *C. chinense,* tandem repeats made up a smaller proportion of the genus (3.8%), and only 0.9% of these repeats were identified as putative satellites. Satellite DNAs are typically species or genus specific (Plohl et al. 2008), and have been found to either be very similar within a species and to vary significantly even between closely related species (Weiss-Schneeweiss et al. 2015). In the present thesis, four clusters of satellite DNAs within each of the two species were further analyzed with DOTTER (Sonnhammer & Durbin 1995), individually and comparatively. *C. annuum* and *C. chinense* were shown to share two satellites. CL81 in *C. annuum* was homologous to CL44 in *C. chinense*, and CL120 in *C. annuum* was homologous to CL101 in *C. chinense* (Fig. 30). Two other satellite DNA repeats were unique in each species. These satellite DNA repeats provide novel tools to analyze the similarities of the two genomes and to identify and analyze the composition of hybrid individuals between the two species.

7. CONCLUSION

The present thesis focused on *Capsicum annuum* and *C. chinense* to provide a cytogenetic characterization of these closely related species, and thus highlight similarities and differences between varieties. By combining data obtained from chromosome counts, flow cytometric measurements, repetitive DNA localization via FISH, and next-generation sequencing, this thesis provides a comprehensive overview of domesticated individuals of *C. annuum* and *C. chinense*.

Overall, the data revealed that the analyzed varieties of *C. annuum* and *C. chinense* have comparable genome size, chromosome size, as well as similar proportions of types of repetitive DNA in their respective genomes.

Chromosomes of the *C. annuum* cultivars were morphologically similar to each other, as are the chromosomes of *C. chinense* cultivars. Between the two species, only minor differences were detected, such as that *C. chinense* varieties generally possessed only one small acrocentric chromosome, while the metaphase plates of most *C. annuum* cultivars showed two. The data also revealed one triploid plant ($2n = 3x = 36$) among 23 diploid individuals ($2n = 2x = 24$), an uncommon occurrence in studies of *Capsicum* so far. Further, the characterization of repetitive DNA fractions within the genomes using NGS-technology showed an abundance of transposable elements, with the largest proportion of elements belonging to *Ty3*-gypsy lineages, and homologous satellites in *C. annuum* and *C. chinense* were identified. It is evident that *C. annuum* and *C. chinense* exhibit significant similarities, but even more so the varieties within each species, which are highly uniform in their chromosomal structure and genomic content. Minor differences, which became apparent upon close examination, are likely associated with changes in the repetitive sequences within the genomes.

In conclusion, the findings of this study not only provide insights into the genomic diversity of the economically significant genus *Capsicum*, but the information collected here may be used in future studies to address questions concerning the mechanisms involved in varietal differentiation or sequences involved in chromosomal and genomic change.

BIBLIOGRAPHY

Aguilar-Meléndez, A., Morrell, P. L., Roose, M. L., Kim, S. C. 2009. Genetic Diversity and Structure in Semiwild and Domesticated Chiles (*Capsicum annuum*; Solanaceae) from Mexico. *American Journal of Botany* 96(6): 1190–1202.

Aguilera, P. M. Debat, H. J., Scaldaferro, M. A., Martí, D. A., Gabriele, M. 2016. FISH-mapping of the 5S rDNA locus in chili peppers (*Capsicum*-Solanaceae). *Anais da Academia Brasileira de Ciências* 88(1): 117–125.

Baranyi, M., Greilhuber, J. 1996. Flow cytometric and Feulgen densitometric analysis of genome size variation in Pisum. *Theoretical and Applied Genetics* 92: 297–307.

Basu, S. K., De, A. K. 2003. Capsicum: historical and botanical perspectives. In: De, A. K., editor. *Capsicum*. London: Taylor & Francis. pp. 1–16.

Benz, B. F. 2001. Archaeological evidence of teosinte domestication from Guilá Naquitz, Oaxaca. *Proceedings of the National Academy of Science in the USA* 98(4): 2104–2106.

Biscotti, M. A., Ettore, O., Heslop-Harrison, J. S. 2015. Repetitive DNA in eukaryotic genomes. *Chromosome Research* 23(3): 415–420.

Boswell, V. R. 1949. Garden pepper, both a vegetable and a condiment. *National Geographics* 96: 166–167.

Brown, C. H., Clement, C. R., Epps, P., Luedeling, E., Wichmann, S. 2013. The Paleobiolinguistics of Domesticated Chili Pepper (*Capsicum* spp.). *Ethnobiology Letters* 4: 1–11.

Bryan, T. M. 2009. Telomeres and the end replication problem. In: Cox, Lynne S., editor. Molecular Themes in DNA Replication. Cambridge: Royal Society of Chemistry. pp. 217–268.

Carrizo García, C., Barfuss, M. H. J., Sehr, E. M., Barboza, G. E., Samuel, R., Moscone, E. A., Ehrendorfer, F. 2016. Phylogenetic relationships, diversification and expansion of chili peppers (*Capsicum,* Solanaceae). *Annals of Botany* 118(1): 35–51.

Carrizo García, C., Sterpetti, M., Volpi, P., Ummarino, M., Saccardo, F. 2013. Wild *capsicums*: identification and in situ analysis of Brazilian species. In: Lanteri, S., Rotino, G. L., editors. Breakthroughs in the genetics and breeding of *Capsicum* and eggplant. Turin: Eucarpia. pp. 205–213.

Carvalho, S. I. C., Ragassi, C. F., Bianchetti, L. B., Reifschneider, F. J. B., Buso, G. S. C., Faleiro, F. G. 2014. Morphological and genetic relationships between wild and domesticated forms of peppers (*Capsicum frutescens* L. and *C. chinense* Jacquin). *Genetics and Molecular Research* 13(3): 7447–7464.

Carvalho, S. I. C., Ragassi, C. F., Bianchetti, L. B., Reifschneider, F. J. B., Buso, G.S. C., Faleiro, F. G. 2014. Morphological and genetic relationships between wild and domesticated forms of peppers (*Capsicum frutescens* L. and *C. chinense* Jacquin). *Genetics and Molecular Research* 13(3): 7447–7464.

Cheema, S. K., Pant, M. R. 2013. Karyotype analysis of seven cultivated varieties of *Capsicum annuum L*. *Caryologia* 66: 70–76.

Chennaveeraiah, M. S., Habib, A. F. 1973. Cytomorphological Studies in a Spontaneous Triploid of *Capsicum annuum L*. *Cytologia* 38: 677–685.

De, A. K. 2003. *Capsicum*. London: Taylor & Francis.

Deschamps, S., Campbell, M. 2010. Utilization of next-generation sequencing platforms in plant genomics and genetic variant discovery. *Molecular Breeding* 25(4): 553–570.

DeWitt, D., Bosland, P. W. 2009. *The Complete Chile Pepper Book*. Portland: Timber Press.
Doebley, J. F., Gaut, B. S., Smith, B. D. 2006. The Molecular Genetics of Crop Domestication. *Cell* 127(7): 1309–1321.
Doyle, J. J., Doyle, J. L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.
Edmonds, J. M. 2012. *Solanaceae*. Richmond, Surrey: Royal Botanic Gardens, Kew.
Egan, A. N., Schlueter, J., Spooner, D. M. 2012. Applications of next-generation sequencing in plant biology. *American Journal of Botany* 99(2): 175–185.
Eshbaugh, W. H. 1993. Peppers: History and Exploitation of a Serendipitous New Crop Discovery. In: Janick, J., Simon, J. E., editors. *New Crops*. New York: John Wiley and Sons. pp. 132–139.
Eyre-Walker, A., Gaut, R. L., Hilton, H., Feldman, D. L., Gaut, B. S. 1998. Investigation of the bottleneck leading to the domestication of maize. *Proceedings of the National Academy of Sciences of the United States of America* 95(8): 4441–4446.
Fukui, K., Nakayama, S. 1996. *Plant chromosomes: laboratory methods*. Boca Raton: CRC Press.
Gailbrath, D. W., Harkins, K. R., Maddox, J. M., Ayres, N. M., Sharma, D. P., Firoozabady, E. 1983. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science* 220: 1049–1051.
Gebhardt, C. 2016. The historical role of species from the Solanaceae plant family in genetic research. *Theoretical and Applied Genet.* 129: 2281–2294.
Greilhuber, J., Borsch, T., Müller, K., Worberg, A., Porembski, S., Barthlott, W. 2008. Smallest Angiosperm Genomes Found in Lentibulariaceae, with Chromosomes of Bacterial Size. *Plant Biology* 8(6): 770–777.
Greilhuber, J., Doležel, J. 2009. 2C or not 2C: a closer look at cell nuclei and their DNA content. *Chromosoma* 118: 391–400.
Griffiths, A. J. F. 2000. *An introduction to genetic analysis.* (7th edition). New York: W.H. Freeman.
Griffiths, A. J. F., Gelbart, W. M., Miller, J. H., Lewontin, R. C. 1999. *Modern Genetic Analysis*. New York: W. H. Freeman.
Gross, B. L., Olsen, K. M. 2010. Genetic perspectives on crop domestication. *Trends in Plant Science* 15: 529–537.
Hammer, K. 1984. Das Domestikationssyndrom. *Kulturpflanze* 32: 11–34.
Hanson, R. E., Islm-Faridi, M. N., Percival, E. A., Crane, C. F., Ji, Y., McKnight, T. D., Stelly, D. M., Price, H. J. 1996. Distribution of 5S and 18-28S rDNA loci in a tetraploid cotton (*Gossypium hirsutum L.*) and its putative diploid ancestors. *Chromosoma* 105(1): 55–61.
Hardigan, M. A., Laimbeer, F., Parker, E., Newton, L., Crisovan, E., Hamilton, J. P., Vaillancourt, B., et al. 2017. Genome diversity of tuber-bearing uncovers complex evolutionary history and targets of domestication in the cultivated potato. *Proceedings of the National Academy of Science in the USA* 114(46): 9999–10008.
Harlan, J. R., Wet, J. M. J., Price, E. G. 1973. Comparative Evolution of Cereals. *Evolution* 27(2): 311–325.
Heiser, C. B., Smith, P. G. 1953. The Cultivated *Capsicum* Peppers. *Economic Botany* 7(3): 214–227.
Heslop-Harrison, J. S. 2000. Comparative genome organization in plants: from sequence and markers to chromatin and chromosomes. *Plant Cell* 12: 617–636.

Heslop-Harrison, J. S., Schwarzacher, T. 2011. Organisation of the plant genome in chromosomes. *Plant J* 66:18–33.

Husband, B.C., Baldwin, S. J., Suda, J. 2013. The incidence of polyploidy in natural plant populations: major patterns and evolutionary processes. In: Leitch, I. J., Greilhuber, J., Doležel, J., Wendel, J. F., editors. *Plant Genome Diversity. Volume 2. Physical Structure, Behavior and Evolution of Plant Genomes.* Wien: Springer Verlag. pp. 209-230.

Huskins, C. L., La-Cour, L. 1930. Chromosome Numbers in *Capsicum*. *The American Naturalist* 64(693): 382–384.

Innan, H., Kim, Y. 2004. Pattern of polymorphism after strong artificial selection in a domestication event. *Proceedings of the National Academy of Science in the USA* 101(29): 10667–10672.

Irish, H. C. 1898. A revision of the genus *Capsicum* with especial reference to garden varieties. *Missouri Botanical Garden Annual Report* 1898: 53–110.

Jang, T.-S., Weiss-Schneeweiss, H. 2015. Formamide-free genomic in situ hybridization allows unambiguous discrimination of highly similar parental genomes in diploid hybrids and allopolyploids. *Cytogenetic and Genome Research* 146(4): 325–331.

Jha, B. T., Dafadar, A., Ghorai, A. 2012. New genetic resource in *Capsicum L.* from Eastern Himalayas. *Plant Genetic Resources* 10(2): 141–144.

Jha, T. B., Saha, P. S., Nath, S., Das, A., Jha, S. 2017. Morphological and cytogenetical characterization of 'Dalle Khursani': a polyploid cultivated *Capsicum* of India. *Scientia Horticulturae* 215: 80–90.

Jiao, Y., Wickett, N. J., Ayyampalayam, S., Chanderbali, A. S., Landherr, L., Raplh, P. E., Tomsho, L. P., Hu, Y., Liang, H., Soltis, P. S., Soltis, D. E., Clifton, S. W., Scharbaum, S. E., Schuster, S. C., Ma, H., Leebens-Mack, J., dePamphilis, C.W. 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* 473: 97–100.

Kantar, M. B., Nashoba, A. R., Anderson, J. E., Blackman, B. K., Rieseberg, L. H. 2017. The genetics and genomics of domestication. *BioScience* 67(11): 971–982.

Kejnovský, E., Hawkins, J., Feschotte, C. 2012. Plant transposable elements: Biology and evolution. In: Leitch, I. J., Greilhuber, J., Doleže, J., Wendel, J. F., editors. *Plant Genome Diversity. Volume 1. Plant genomes, their residents, and their evolutionary dynamics*. Wien: Springer Verlag. pp. 17–34.

Kim, S., Park, M., Yeom, S., Kim, Y., Lee, J. M., Lee, H., et al. 2014. Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. *Nature genetics* 46(3): 270–279.

Kraft, K. H., Brown, C. H., Nabhan, G. P., Luedeling, E., De Jesús Luna Ruiz, J., D'Eeckenbrugge, G. C., Hijmans, R. J., Gepts, P. 2014. Multiple lines of evidence for the origin of domesticated chili pepper, *Capsicum annuum*, in Mexico. *Proceedings of the National Academy of Science in the USA* 111(17): 6165–6170.

Kulkarni, M., Borse, T. 2010. Induced polyploidy with gigas expression for root traits in *Capsicum annuum (L.). Plant Breeding* 129(4): 461–464.

Kwon, J., Kim, B. 2009. Localization of 5S and 25S rRNA Genes on Somatic and Meiotic Chromosomes in *Capsicum* Species of Chili Pepper. *Molecules and Cells* 27: 205–209.

Leitch, A. R., Leitch, I. J. 2008. Genomic plasticity and the diversity of polyploid plants. *Science* 320: 481–483.

Leitch, I. J., Bennett, M. D. 2004. Genome downsizing in polyploid plants. *Biological Journal of the Linnean Society* 82: 651–663.

Leitch, I. J., Greilhuber, J., Doleže, J., Wendel, J. F. 2013. *Plant Genome Diversity. Volume 2. Physical Structure, Behavior and Evolution of Plant Genomes.* Wien: Springer Verlag.

Levin, D. A. 2002. *The Role of Chromosomal Change in Plant Evolution*. New York: Oxford University Press.

Levitzky, G. A. 1931. The karyotype in systematics. *Bulletin of Applied Botany, of Genetics and Plant Breeding* 27: 220–240.

Lippert, L. F., Smith, P. G., Bergh, B. O. 1966. Cytogenetics of the Vegetable Crops. Garden Pepper, *Capsicum* Sp. *Botanical Review* 32(1): 24–55.

Luceño, M., Guerra, M. 1996. Numerical variations in species exhibiting holocentric chromosomes: a nomenclatural proposal. *Caryologica* 49: 301–309.

Lynch, M. 2007. *The Origins of Genome Architecture*. Sunderland, Mass.: Sinauer Associates.

Lysák, M. A., Schubert, I. 2013. Mechanisms of chromosome rearrangements. In: Leitch, I. J., Greilhuber, J., Doležel, J., Wendel, J. F., editors. *Genome Diversity. Volume 2. Physical Structure, Behavior and Evolution of Plant Genomes*. Wien: Springer Verlag. pp. 137-147.

Mehrotra, S., Goyal, V. 2014. Repetitive sequences in plant nuclear DNA: Types, distribution, evolution and function. *Genomics, Proteomics & Bioinformatics* 12(4): 164–171.

Meyer, R. S., Purugganan, M. D. 2013. Evolution of crop species: genetics of domestication and diversification. *Nature Reviews Genetics* 14(12): 840–852.

Moscone, E. A., Baranyi, M., Ebert, I., Greilhuber, J., Ehrendorfer, F., Hunziker, A. T. 2003. Analysis of Nuclear DNA Content in *Capsicum* (Solanaceae) by Flow Cytometry and Feulgen Densitometry. *Annals of Botany* 92: 21–29.

Moscone, E. A., Loidl, J., Ehrendorfer, F., Hunziker, A. T. 1995. Analysis of active nucleolus organizing regions in *Capsicum* (Solanaceae) by silver staining. *American Journal of Botany* 82: 276–287.

Moscone, E. A., Scaldaferro, M. A., Gabriele, M., Cecchinni, N. M., Ehrendorfer, F. 2007. The evolution of chili peppers (*Capsicum* – Solanaceae): a cytogenetic perspective. *Acta Horticulturae* 745(2): 137–169.

Novák, P., Neumann, P., Macas, J. 2010. Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. *BMC Bioinformatics* 11: 378.

Novák, P., Neumann, P., Pech, J., Steinhaisl, J., Macas, J. 2013. RepeatExplorer: a galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* 29(6): 792–793.

Nunez, R. 2001. DNA Measurement and Cell Cycle Analysis by Flow Cytometry. *Current Issues in Molecular Biology* 3(3): 67–70.

Ormerod, M. G. 2009. *Flow Cytometry: A basic introduction*. http://flowbook.denovosoftware.com/ (5 Aug. 2018).

Park, M., Park, J., Kim, S., Kwon, J.-K., Park, H. M., Bae, I. H., Yang, T.-J., Lee, Y.-H., Kang, B.-C., Choi, D. 2012. Evolution of the large genome in *Capsicum annuum* occurred through accumulation of single-type long terminal repeat retrotransposons and their derivatives. *The Plant Journal* 69: 1018–1029.

Pellicer, J., Fay, M. F., Leitch, I. J. 2010. The largest eukaryotic genome of them all?. *Botanical Journal of the Linnean Society* 164(1): 10–15.

Perry, L., Dickau, R., Zarrillo, S., Holst, I., Pearsall, D. M., Piperno, D. R., et al. 2007. Starch fossils and the domestication and dispersal of chili peppers (*Capsicum spp. L.*) in the Americas. *Science* 315(5814): 986–988.

Pickersgill, B. 1971. Relationships between weedy and cultivated forms in some species of chili peppers (genus *Capsicum*). *Evolution* 25: 683–691.

Pickersgill, B. 1997. Genetic resources and breeding of *Capsicum* spp. *Euphytica* 96(1): 129–133.

Pickersgill, B. 2007. Domestication of Plants in the Americas: Insights from Mendelian and Molecular Genetics. *Annals of Botany* 100: 925–940.

Pickersgill, B., Heiser, C. B. 1976. Cytogenetics and evolutionary change under domestication. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 275(936): 55–69.

Plohl, M., Luchetti, A., Mestrović, N., Mantovani, B. 2008. Satellite DNAs between selfishness and functionality: Structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene* 409: 72–82.

Purugganan, M. D., Fuller, D. Q. 2009. The nature of selection during plant domestication. *Nature* 457: 843–848.

Raghavan, T., Venkatasubban, K. 1949. Studies of the South Indian Chillies. *Proceedings of the Indian Academy of Sciences* 12(2): 29–46.

Ramsey, J., Schemske, D. W. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* 291(1): 467–501.

Rieseberg. L. H. 2001. Chromosomal rearrangements and speciation. *Trends in Ecology and Evolution* 16: 351–358.

Romero-da Cruz, M. V., Urdampilleta, J. D., Forni Martins, E. R., Moscone, E. A. 2017. Cytogenetic markers for the characterization of *Capsicum annuum* L. cultivars. *Plant Biosystems* 151(1): 84–91.

Salmon, A., Ainouche, M. 2015. Next-generation sequencing and the challenge of deciphering evolution of recent and highly polyploid genomes. In: Hörandl, E., Appelhans, M. S., editors. *Next-generation sequencing in plant systematics*. Königstein: Koeltz Scientific Books. pp. 93-118.

Scaldaferro, M. A., Romero-da Cruz, M. V., Cecchini, N., Moscone, E. A. 2015. FISH and AgNor-mapping of the 45S and 5S rRNA genes in wild and cultivated *Capsicum* (Solanaceae). *Genome* 59(2): 95–113.

Scaldaferro, M. A., Seijo, J. G., Acosta, M. C., Barboza, G. E., Ducasse, D. A., Moscone, E. A. 2006. Genomic characterization of the germplasm in peppers (*Capsicum* – Solanaceae) by fluorescent *in situ* hybridization. *Plant Science* 43: 291–297.

Schmidt, T., Heslop-Harrison, J. S. 1998. Genomes, genes and junk: the large-scale organization of plant chromosomes. *Trends in Plant Science* 3: 195–199.

Schubert, I. 2007. Chromosome evolution. *Current Opinion in Plant Biology* 10: 109–115.

Schwarzacher, T., Heslop-Harrison, P. 2000. *Practical in situ hybridization*. Oxford: Bios Scientific Publishers.

Sharma, S. K., Vij, A. S., Sharma, M. 2013. Mechanisms and clinical uses of capsaicin. *European Journal of Pharmacology* 720: 55–62.

Smith, P. G., Heiser, C. B. 1951. Taxonomic and Genetic Studies on the Cultivated Peppers, *Capsicum annuum L.* and *C. frutescens L.*". *American Journal of Botany* 38(5): 362–368.

Smith, P. G., Heiser, C. B. 1957. Taxonomy of *Capsicum sinense* Jacq. and the Geographic Distribution of the Cultivated *Capsicum* Species. *Bulletin of the Torrey Botanical Club* 84(6): 413–420.

Soltis, D. E., Albert, V. A., Leebens-Mack, J., Bell, C. D., Paterson, A. H., Zheng, C., Sankoff, D., dePamphilis, C. W., Wall, P. K., Solits, P. S. 2009. Polyploidy and angiosperm diversification. *American Journal of Botany* 96: 336–348.

Soltis, P. S., Marchant, D. B., Van der Peer, Y., Soltis, D. S. 2015. Polyploidy and genome evolution in plants. *Current Opinion in Genetics & Development* 35: 119–125.

Sousa, W. R. d. N., Lopes, A. C. d. A., de Carvalho, R., Gomes, R. L. F., Peron, A. P. 2015. Karyotypic characterization of *Capsicum* sp. accessions. *Acta Scientiarum: Agronomy* 37(2): 147–153.

Stace, C. A. 2000. Cytology and cytogenetics as fundamental taxonomic resources for the 20th and 21st centuries. *Taxon* 49: 451–477.

Stebbins, G. L. 1971. *Chromosomal Evolution in Higher Plants*. London: Arnold.

Sturtevant, E. 1885. Kitchen garden esculents of American origin. II. Peppers. *American Naturalist* 19: 544–550.

Tang, H., Bowers, J. E., Wang, X., Ming, R., Alam, M., Paterson, A. H. 2008. Synteny and Collinearity in Plant Genomes. *Science* 320(5875): 486–488.

Tang, H., Sezen, U., Paterson, A. H. 2010. Domestication and plant genomes. *Current Opinion in Plant Biology* 13: 160–166.

Temsch, E. M. 2003. Genome Size Variation in Plants with Special Reference to the Genus *Arachis* (Fabaceae), and the Methodological Improvement by Application of a new Optica Immersion-Gel in Flow Cytometry. PhD thesis, Department of Systematic and Evolutionary Botany, University of Vienna.

Temsch, E. M., Greilhuber, J., Krisai, R. 2010. Genome size in liverworts. *Preslia* 82: 63–80.

Tenaillon, M. I., Hollister, J. D., Gaut, B. S. 2010. A triptych of the evolution of plant transposable elements. *Trends in Plant Science* 15: 471–478.

Weiss-Schneeweiss, H., Leitch, A. R., McCann, J., Jang, T.-S., Macas, J. 2015. Employing next generation sequencing to explore the repeat landscape of the plant genome. In: Hörandl, E., Appelhans, M., editors. *Next Generation Sequencing in Plant Systematics*. Königstein: Koeltz Scientific Books. pp. 155-180.

Weiss-Schneeweiss, H., Schneeweiss, G. M. 2013. Karyotype diversity and evolutionary trends in angiosperms. In: Leitch, I. J., Greilhuber, J., Doležel, J., Wendel, J. F., editors. *Plant Genome Diversity. Volume 2. Physical Structure, Behavior and Evolution of Plant Genomes.* Wien: Springer Verlag. pp. 209-230.

Weiss-Schneeweiss, H., Stuessy, T. F., Siljak-Yakovlev, S., Baeza, C. M., Parker, J. 2003. Karyotype evolution in South American species of *Hypochaeris* (Asteraceae, Lactuceae). *Plant Systematics and Evolution* 241: 171–184.

Weiss-Schneeweiss, H., Tremetsberger, K., Schneeweiss, G. M., Parker, J. S., Stuessy, T. F. 2008. Karyotype Diversification and Evolution in Diploid and Polyploid South American *Hypochaeris* (Asteraceae) Inferred from rDNA Localization and Genetic Fingerprint Data. *Annals of Botany* 101: 909–918.

Wright, S. I., Bi, I. V., Schroeder, S. G., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S. 2005. The Effects of Artificial Selection on the Maize Genome. *Science* 308: 1310–1314.

Zhang, X., Liu, G., Yan, L., Zhao, Y., Chang, R., Wu, L. 2003. Creating triploid germplasm via induced 2n pollen in *Capsicum annuum L. The Journal of Horticultural Science and Biotechnology* 78(1): 84–88.

ZUSAMMENFASSUNG

Die vorliegende Arbeit umfasst eine Genomanalyse ausgewählter Vertreter der Gattung *Capsicum*. Die Gattung gehört zur Familie der Nachtschattengewächse und ist vor allem wegen ihrer Früchte von weltweiter kultureller und wirtschaftlicher Bedeutung. Nach dem derzeitigen Stand der Forschung beinhaltet *Capsicum* 35 Arten, von denen fünf als domestiziert gelten. Die im Rahmen dieser Diplomarbeit analysierten Pflanzen sind den Spezies *Capsicum annuum* und *Capsicum chinense* zugeordnet.

Der erste Teil der Arbeit beinhaltet eine Literaturanalyse und beschäftigt sich einerseits mit grundlegenden Prozessen der Veränderung pflanzlicher Genome, andererseits mit der Charakterisierung der Gattung *Capsicum*. Der praktische Teil der Arbeit befasst sich mit der Beschreibung 25 ausgewählter Indivuduen von *C. annuum* und *C. chinense* und dem Vergleich der beiden Spezies.

Chromosomenzählungen mittels Feulgen-Färbung bestätigten die in der Literatur angegebenen Werte für diploide Organismen ($2n = 2x = 24$). Genomgrößen wurden mittels Durchflusszytometrie gemessen und ergaben Werte zwischen 3.4081 und 3.5796 pg, welche ebenfalls im Rahmen der bisher gesammelten Ergebnisse für die beiden Spezies liegen. Bei einem Individuum, *C. chinense* cv. ‚Japonaise', wurden ein deutlich höherer DNA-Gehalt des Chromosomensatzes gemessen (5.2253 pg). Chromosomenzählungen bestätigten, dass es sich um ein triploides Individuum hält ($2n = 3x = 36$). Polyploidie wurde in domestizierten Vertretern der Gattung *Capsicum* bisher selten beschrieben. Die Analyse beinhaltete weiters den Nachweis von ribosomaler DNA durch Fluoreszenz-*in-situ*-Hybridisierung ausgewählter Individuen. Schließlich wurden in der vorliegenden Arbeit die Genome von *C. annuum* und *C. chinense* mittels „Next-generation Sequencing" analysiert, und dadurch Anteile repetitiver DNA-Elemente in den Genomen der beiden Arten identifiziert. Die Ergebnisse der vergleichenden Untersuchungen wurden in einem phylogenetischen Kontext interpretiert. Die hohe Ähnlichkeit der beiden Arten im Hinblick auf ihre Chromosomenzahl, Genomgröße und Proportionen von repetitiver DNA ist ihrer nahen Verwandtschaft zuzuschreiben.