DEPARTAMENTO DE CIÊNCIAS DA VIDA

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Incidence of polyploidy and genome evolution in Scrophulariaceae *s.l.*

Digitalis mariana* subsp. *heywoodii

Mariana Oliveira e Castro

2011

DEPARTAMENTO DE CIÊNCIAS DA VIDA

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Incidence of polyploidy and genome evolution in Scrophulariaceae *s.l.*

Dissertação apresentada à Universidade de Coimbra para cumprimento dos requisitos necessários à obtenção do grau de Mestre em 2011, realizada sob a orientação científica do Professor Doutor João Carlos Mano Castro Loureiro (Universidade de Coimbra) e da Doutora Sílvia Raquel Cardoso Castro Loureiro (Universidade de Coimbra)

Mariana Oliveira e Castro

2011

Acknowledgements

First, I want to thank to all the people who always supported me and directly or indirectly contributed for the accomplishment of the present work.

I am grateful to João Loureiro and Sílvia Castro for accepting the supervision of this work, for all their support and friendship and for the continuous encourage for making science.

I also thank to Dr Jorge Paiva for taxonomic lessons and for all his support in the corrent identification of some plant material.

I want to thank to the *Index Semina* of COI, UPT, ISAUTL and MHNM for sending Scrophulariaceae seeds.

I want to thank to Prof. Conceição Santos and Helena Oliveira for enabling the use of the flow cytometer of the Laboratory of Biotechnology and Cytomics (Department of Biology, University of Aveiro) in the first stages of this thesis.

I do not forget all the members (official or borrowed) of the Plant Ecology and Evolution Group for their friendship and for the excellent work environment.

Also, I am grateful to all the people of the Department of Life Sciences that somehow helped me during the development of this thesis, in special Manuela Patão, Ludovina Lopes, Arménio Matos, Luísa Ramalho and Pedro Cunha.

A special thanks also to my friends for all their support and encouragement, in special, Ana Maranha, Daniela Luís and Laura Perpétuo.

Finally, but not least important, this thesis is dedicated to my parents, brother and grandmothers, who continuously supported me and had a inexhaustible patience and understanding of all my choices, to my beloved José Ricardo, and in special to my uncle Domingos, who I miss.

Table of Contents

i. Abbreviations

AAl – Alto Alentejo

Ag – Algarve

B – *Bellis perenis*

BA – Beira Alta

BAl – Baixo Alentejo

BB – Beira Baixa

BL – Beira Litoral

COI – Universidade de Coimbra

CV – coefficient variation

DL – Douro Litoral

E – Estremadura

e.g. – (L. *exempli gratia*) for example

EDTA $Na_2.2H_2O$ – Ethylenediaminetetraacetic acid

et al. – (L. *et alia*) and other

FISH – fluorescence in situ hybridization

FL – fluorescence pulse integral

FS – forward light scatter

G – *Glycine Max*

i.e. – (L. *id est*) that is

IS – Index seminum

ISAUTL – Instituto Superior de Agronomia Universidade Técnica de Lisboa

JC – Joana Costa

JL – João Loureiro

JP – Jorge Paiva

Mbp – mega base pair

MC – Mariana Castro

$MgCl_2.6H_2O$ – Magnesium Chloride Hexahydrate

Mi – Minho

MNHN – Museu Nacional de História Natural

n – number of

NaCl - Sodium chloride

P – *Pisum sativum*

pg – pictograms

POP – natural populations

PVP – Polyvinylpyrrolidone 10

R – *Raphanus sativus*

R – Ribatejo

R^2 – coefficient

S – *Solanum lycopersicum*

s.d. – standard deviation

s.l. – (L. *sensu lato*) in the broad sense

s.s. –- (L. *sensu sctricto*) in the sense of

SC – Sílvia Castro

SS – side light scatter

subsp – subspecie

TM – Trás-os-Montes

Tris.HCl – Tris(hydroxymethyl)aminomethane

UPT – Universidade Portucalensis

WPB - woody plant buffer

Z – Zea mays

δ – degrees of freedom

ii. Resumo

Na última década, o recurso a marcadores moleculares influenciou fortemente a filogenia actual das Angiospérmicas, com diversos *taxa* a transitarem para novas famílias ou para outras famílias já existentes. O conteúdo em ADN nuclear contribuiu para esta discussão, sendo considerado um caracter importante na biossistemática e na ecologia e biologia das populações. Também, como novas entidades evolutivas podem surgir num único evento genético, a poliploidia foi proposta como um mecanismo importante na génese de biodiversidade. Nos últimos anos, tem existido um interesse crescente em estudos relacionados com a evolução do tamanho do genoma e com a incidência da poliploidia, apesar de na família das Scrophulariaceae, existirem muito poucos estudos disponíveis. Para além disto, e contrariamente ao observado noutras regiões do globo, existem poucos estudos focados na incidêncida da poliploidia na flora da Península Ibérica, um centro importante de diversificação. Tendo estes factos em consideração, os objectivos principais desta tese de Mestrado foram avaliar a importância do tamanho do genoma, especialmente como marcador taxonómico, e o papel da poliploidia como um processo de génese e manutenção da diversidade das Scrophulariaceae *s.l.* na Península Ibérica. Para o efeito, foram realizadas análises em larga escala da variação do tamanho do genoma e do nível de ploidia usando a citometria de fluxo ao longo da região oeste da Península Ibérica. Cento e sessenta e duas populações de 59 *taxa* distintos foram analisadas, tendo sido analisados 3 indivíduos por população para estimativas do tamanho do genoma e 30 indivíduos para análises do nível de ploidia. Dos 59 *taxa* analisados, 86% representam as primeiras estimativas do tamanho do genoma. A maioria das espécies de Scrophulariaceae apresenta um tamanho do genoma muito pequeno ou pequeno (2C ≤ 7.0 pg), com uma espécie apenas a apresentar um tamanho do genoma intermédio. Também, na maioria dos géneros analisados, foi possível utilizar este caracter para separar alguns dos *taxa*, independentemente de estes géneros serem homoplóides (e.g., *Digitalis*, *Linaria*) ou heteroploides (e.g., *Veronica*). Ainda, foram observados outros fenómenos relacionados com a evolução do genoma, em particular, variação intra-específica do tamanho do genoma em alguns géneros (e.g., *Scrophularia*) e disploidia em *Verbascum*. No que diz respeito à poliploidia, apesar de terem sido detectados novos níveis de ploidia em *Veronica*, não foi descoberto nenhum *taxa* com citotipos múltiplos. Estes dados revelam que aparentemente, a poliploidia não é um dos principais mecanismos de especiação nas Scrophulariaceae, pelo menos nesta região.

Palavras-chave: biossistemática, citometria de fluxo, evolução do genoma, Península Ibérica, poliploidia, Scrophulariaceae, tamanho do genoma, taxonomia.

iii. Abstract

In the last decade the genomic studies using DNA markers have strongly influenced the current phylogeny of Angiosperms, with several plant *taxa* being moved to new or existing families. The nuclear DNA content itself has contributed to this discussion being considered an important character in biosystematics and more recently in ecology and population biology. Also, as new evolutionary entities can arise in a single genetic event, polyploidy has been proposed as an important mechanism for generating biodiversity. In recent years, there has been an increasing focus in studies related with genome size evolution and polyploidy incidence, despite that in Scrophulariaceae only a few works are available. Furthermore, contrarily to other regions, only a few studies concerning polyploidy incidence were focused on the Iberian Peninsula flora, an important center of diversification. Considering this, the main objectives of this Master thesis were to assess the value of genome size, mostly as a taxonomic marker, and the role of polyploidy as a process of genesis and maintenance of plant diversity in Scrophulariaceae *s.l.* in Iberian Peninsula. For that, large scale analyses of genome size and ploidy level variation were performed using flow cytometry across the Western Iberian Peninsula. One hundred sixty two populations of 59 distinct *taxa* were analyzed, with 3 individuals per population being analysed for genome size and 30 individuals for ploidy level. From the 59 sampled *taxa*, 86% were first estimates of genome size. The majority of the Scrophulariaceae species presented very small to small genome sizes (2C ≤ 7.0 pg), with only one species presenting an intermediate genome size. Furthermore, in most of the analysed genera it was possible to use this character to separate several *taxa*, independently if these genera were homoploid (e.g., *Digitalis*, *Linaria*) or heteroploid (e.g., *Veronica*). Also, other genome related phenomena were detected, as intraspecific variation of genome size in some genera (e.g., *Scrophularia*) and disploidy in *Verbascum*. With respect to polyploidy, despite a few new DNA ploidy levels have been detected in *Veronica*, no multiple cytotypes have been found in any *taxa*, revealing that polyploidy apparently is not among the main mechanisms of speciation in Scrophulariaceae, at least in this region.

Key words: biosystematics, flow cytometry, genome evolution, genome size, Iberian Peninsula, polyploidy, Scrophulariaceae, taxonomy.

Chapter 1
INTRODUCTION

As a consequence of globalization, continuous growth of human population, resource exploration and climate changes (Pimm et al. 1995, Novacek and Cleland 2001, Warren et al. 2001, Brook et al. 2003), the degradation of natural habitats is reaching unprecedented proportions in history, and has been resulting in massive extinctions of biota (Sala et al. 2000, Novacek and Cleland 2001). For example. about 40-50% of the ice-free land surface was transformed by human activity, using one-third of the terrestrial net primary productivity and causing the extinction of 5-20% of the species in many groups of organisms (Brook et al. 2003, Chapin 2003). Nowadays, the rates are 100 to 1,000 higher than in the pre-human levels (Pimm et al. 1995), with about 25% of the 250,000 estimated species of vascular plants potentially becoming extinct within the next 50 years (Raven 1987). Also, dramatic declines of biodiversity led to reductions in the number of of species and subspecies, races and populations (Myers 1989). Still, to counterbalance this scenario, there are still several reports of the emergence of new entities, despite at much lower rates. For a correct evaluation of this phenomenon it is very important to develop strong efforts to investigate and understand the mechanisms of species neogenesis. In particular, efforts should be focused in studying groups in need of taxonomic revision, which may result in the re-definition of species boundaries (Greilhuber and Speta 1985), in the detection of undescribed *taxa* (Maxte et al. 1991), or in the revelation that a species has been misidentified (Yeater et al. 2004).

In the last decade the study of genome using DNA markers has strongly influenced the current phylogeny of Angiosperms, with several plant groups being moved to new or existing orders/families (Stevens 2011 onwards). The nuclear DNA content itself has contributed to this discussion being considered an important character in biosystematics and more recently in ecology and population biology (Loureiro 2007).

1.1. Nuclear DNA content and its significance

The knowledge of the genome has been increasingly important in many areas of plant research, including taxonomy and biosystematics, ecology and population biology. Genomes represent a distinct and legitimate level of organization, with unique and own evolutionary histories. Genome size is one of its intrinsic characteristics, being considered a constant species-specific character that can help to explain relationships between species (Gregory 2001). However, before focusing on the significance of genome size, it is important to explore the current terminology and nomenclature of “C-value”.

The term “C-value”, in its short history, has suffered from a misleading nomenclature and significance and only recently with the work of Greilhuber and co-authors a standardized nomenclature was achieved (Greilhuber et al. 2005). The first meaning of “C-value” was assigned by Swift in 1950 and was relative to “constant” DNA content, i.e., the amount of DNA that was characteristic of a particular genotype (Smith 1950). In 1976, Bennett and Smith defined it as the “DNA content of one unreplicated haploid chromosome complement” (Bennett and Smith 1976). Later, the term “C-value” was associated with the DNA content of the complete chromosome complement and the term “genome size” was restricted to the DNA content of the monoploid chromosome set (Bennett et al. 1998). Recently the terminology for genome size was standardized (Greilhuber et al. 2005), with the adjectives “holoploid” (C-value) and “monoploid” (Cx-value) being introduced to distinguish between DNA content of the unreplicated haploid genome (with chromosome number *n*) and DNA content of a single chromosome set of an organism (with chromosome number *x*), respectively (Greilhuber et al. 2005). In summary, C-value refers to half the somatic DNA content (the 2C-value), and Cx-value is the 2C-value divided by the ploidy level (i.e., the number of copies of the genome). In diploid organisms, C-value and Cx-value are equivalent, but not in polyploids. For these, the haploid state contains more than a single chromosome set.

Genome sizes can be expressed in mass units (picograms, pg), or in number of base pairs (bp). The conversion for both units is possible using the following equations (Doležel et al. 2003):

$$\text{DNA content (bp)} = (0.978 \times 10^9) \times \text{DNA content (pg)}$$

$$\text{DNA content (pg)} = \text{DNA content (bp)} / (0.978 \times 10^9)$$

The first nuclear DNA content estimate in plants dates to the 1950’s; however only in the 1980’s the number of estimates started to increase at good rates. This is related with the advancement of flow cytometry, a rapid, robust and reliable technique that can be used for this purpose (see section 1.3 for details). In April of 1997, the first electronic version of the “Angiosperm DNA C-values database” was launched (Bennett and Leitch 2010). At the present, this database covers all the main plant groups and includes genome size estimates for up to 7,058 species (http://data.kew.org/cvalues/). The present knowledge of genome size in plants is summarized in Table 1.

Table 1. Descriptive statistics (min., minimum; max., maximum; mean and range) of 1C DNA values in the major plant groups, together with the level of species representation of C-values data using the latest release of the Plant DNA C-values database (Bennett and Leitch 2010).

Plants group	Min. (pg)	Max. (pg)	Mean (pg)	Range (max./min.)	No. species with DNA C-values	No. species recognized	Species representation (%)
Bryophytes	0.17	7.97	0.66	47	232	≈ 18,000	≈ 1.3
Pteridophytes	0.09	72.68	11.77	808	82	≈ 11,900	≈ 0.7
Gymnosperms	2.25	36.00	18.57	16	204	≈ 730	≈ 27.9
Angiosperms	0.06	152.20	5.94	2,537	6287	≈250,000	≈ 2.5

Of all plants groups, Angiosperms present the most remarkable variation in holoploid genome size, spanning nearly a 2,500-fold range, with *Genlisea margaretae* (Lentibulariaceae, 1C = 0.06 pg; Greilhuber et al. 2006) and *Paris japonica* (Melanthuaceae, 1C = 152.20 pg; Pellicer et al. 2010) presenting the smallest and the largest genomes discovered so far. In this plant group only approximately 2.5% (6,287) of the recognized species have their genome size estimated (Bennett and Leitch 2010).

Despite the small representation of estimates it is already possible to find a large variation in genome size among different taxonomic groups. This highlights the relevance of genome size as a taxonomic and/or ecological marker in particular plant groups. Also, this variation in amount of DNA content (or lack thereof) has been a central focus on evolutionary biology, an one important tool to know the structure of genetic information, its evolution and function, and understand the biological basis of the diversity and its adaptive value in ecological, evolutionary and taxonomic interpretations (Gregory 2005b, Greilhuber et al. 2010).

At the beginning of this research area, variations in genome size were seen as rather enigmatic due to the lack of correlation between the amount of nuclear DNA content and organism's complexity. This was named the "C-value paradox". Later, this paradox was solved by the discovery that much of the variation was due to repetitive non-codified DNA and, therefore renamed the "C-value enigma" (Gregory 2001, Gregory 2005a). Indeed, genome evolution is now considered to be a highly dynamic and bidirectional process and its size result from a dynamic balance between expansion and contraction forces (increasing and decreasing, respectively) (Bennett and Leitch 2005). Generally, polyploidy (explored in section 1.2) is one of the mechanisms that may lead to increases in genome size. In homoploid plants (plants

with the same number of chromosomes but different genome sizes), genome expansion is due to amplification and insertion of transposable genetic elements (different amounts of noncoding, repetitive DNA; Vitte and Bennetzen 2006) and evolution and amplification of satellite repeats (variation in the number and the proportion of minisatellites and microsatellites; Lim et al. 2006). Environment conditions may have an important role in genome evolution, as it may modulate the transcriptional activity of (retro)transposons (Kalendar et al. 2000). Relatively to the loss of genome size, it is associated with deletional mechanisms like unequal intra-strand homologous recombination, illegitimate recombination and/or higher rate of nucleotide deletion over insertion (Bennetzen et al. 2005).

Despite being considered a not so common phenomenon, some species may present variation in the amount of DNA among individuals within and/or between populations, i.e., intraspecific variation of genome size. This variation can be attributed to chromosomal differences (aneuploidy, polyploidy, supernumerary B-chromosomes, sex chromosomes) and cryptic species (Greilhuber 1998), and to polymorphisms in A chromosomes (heterochromatic knobs and differential deletion of transposable element remnants) (Gregory 2005a). Studies of intraspecific variation of genome size usually require the detection of small differences in the amount of DNA, and thus are technologically more demanding (high precision and resolution are usually needed). Once again, if methodological best practices are undertaken, flow cytometry is the ideal technique to accomplish this aim.

In biosystematics, ecology and evolution, genome size usage can be summarized in three main goals: as a taxon-specific marker in biosystematics, to predict the correlation of genomes size with several phenothypic, physiological and/or ecological characteristics (the nucleotypic effect) and to understand the dynamics of genome evolution (studying inter- and intraspecific variation “patterns” in genome size). Some of these goals are described in more detail below:

1.1.1 Importance of genome size in biosystematics

Despite in many cases, taxonomists relied on chromosomal data as an important marker in organisms division (Stace 2000, Ekrt et al. 2009), only recently genome size was been taken into consideration (Kron et al. 2007). As referred above, with a few exceptions, the variation of genome size is mostly constant among species with a larger variation observed among higher taxonomic categories, with some evolutionary relationships (Greilhuber 1998, 2005). So, genome size can be used a supportive tool

to discriminate *taxa* and resolve complex low-level taxonomies. In homoploid plants, genome size has high value to distinguish groups with phenotypic similarities, with a low number of distinct morphological characters, with continuous morphological variations and/or groups prone to inter-specific hybridization or with complex evolutionary histories (e.g., allopolyploids) (Loureiro et al. 2010).

For example, in *Helleborus* (one of the first FCM studies on homoploid groups), all species have 32 chromosomes in their somatic cells, but it was found that their genome is highly variable (Zonneveld et al. 2001). In this case, it was possible distinguish several species and the variation in genome size corresponds to the sectional division. After this study, many others focused in assessing the variation in genome size at species level were published, contributing to an increase in the number of genome size estimates. Another example that proves the value of using genome size as taxonomical marker is the study of Iberian *Festuca* species, in particular the distinction of two subspecies of *Festuca ampla*, subsps. *ampla* and *transtagana*, according to their genome size (Loureiro et al. 2007a).

The identification of homoploid hybrids through genome size is another useful application in biosystematics. Generally, it is supposed that hybrids present a genome size intermediate and non-overlapping with that of parental species (Trucco et al. 2006). In case the genome of parental species differ sufficiently (by at least 7%), it is possible to detect homoploid hybrids using flow cytometry. This was the case of *Elytrigia repens* and *E. intermedia* and the subsequent hybrid (Mahelka et al. 2005). In *Cirsium* a high incidence of inter-specific hybridization was detected. Furthermore, a negative relationship between genome size and the incidence of hybrids was observed, *i.e.*, species with smaller genome size had a higher hybridization frequency (Bureš et al 2004).

In some plants groups, such as allopolyploids (explained in section 1.3) with genome size differences in parental species and similar phenology, it may be possible to deduce the evolutionary relationships and genome constitution of hybrid species (Suda et al. 2007a). It is theoretically expected that, in polyploids, the genome of the new entity has the sum of genome size of the progenitors. However, in practice this process is usually accompanied by losses and/or gains in DNA. In 2008, Leitch and co-authors, observed a high difference between actual (determined by FCM) and expected (sum of genome size of parental species or most closely related to the diploid progenitor) genome size in allopolyploid *Nicotiana* species, revealing genome downsizing in some polyploids and an increase in others, with the former being more frequent. Differences between the actual and expect values were, in general, positively

correlated with evolutionary age (Leitch et al. 2008). The authors hypothesized that the observed DNA loss could be a selection mechanism to minimize genetic instability or the phenotypic effects of having an increased nucleus and cell size.

1.1.2 Biological correlates of genome size

The enormous genome size variation raises many questions regarding biological correlates, stability and plasticity, function and effect, and selective significance and inertness of this character (Greilhuber et al. 2010). In 1971, M. D. Bennett proposed that DNA amount of an organism could influence its phenotype, through the genotype (information contained in genome) and through the nucleotype (mass and volume of the genome), which could impose drastic limits on the range of phenotypes expressed by genic control (the so-called “nucleotypic theory”) (Bennett 1971). However, only recently this character has been integrated into ecological predictions and models. Some of the most interesting predictions include the positive relationships between genome size and cell size (Gregory 2005b), duration of cell division (Bennett 1977), guard cell length and size of epidermal cells (Beaulieu et al. 2008), pollen volume (Bennett 1972), seed size (Bennett 1972) and seed mass (Beaulieu et al. 2007). In the case of stomatal density, Beaulieu and co-authors observed a negatively relationship with genome size and as stomata size and frequency influence carbon fixation and efficient uses of water, variations in genome size may alter the plant physiology. Also, these observations may be useful to predict the natural distribution of species (*e.g.*, species with large genome sizes have a low frequency of small stomata, which may be related with adaptation to dry environments; Beaulieu et al. 2008).

Another possible application of the study of genome size is the distinction, in early developmental stages of, male and female individuals of dioecious plants with heteromorphic chromosomes (*e.g.*, *Cannabis*, *Coccinia*, *Silene* and *Viscum*), if there are significant differences in the size of sexual chromosomes (Loureiro et al. 2010). However in *Silene latifolia*, besides sex related differences, some intraspecific genome size variation was detected. This variation in genome size was reflected phenotypically in the calyx diameter, an adaptive characteristic in this dioecious species (Meagher and Costich 1994).

In terms of broader evolutionary implications, it is generally assumed that genera with larger genomes are less speciose, suggesting that having a larger amount of DNA can be disadvantageous (Knight et al. 2005). Very small genomes clearly prevail in species-rich genera that underwent adaptive radiation, providing support for a

relationship between the content of nuclear DNA and bursts of speciation, especially in insular populations (Loureiro at al. 2010).

1.2. The role of polyploidy in plant diversity

Plant speciation is characterized by the evolution of barriers to gene exchange between populations that previously had the possibility to mate freely. Reproductive isolation can be achieved through a combination of several barriers from different spatial distributions due to dissimilar habitat preferences, to different phenologies, different pollinator guilds, pre- or post- fertilization incompatibilities and low viability of the offspring (hybrids) (Castro et al. 2011). Because of immediate shifts in plant morphology and ecological tolerances of polyploid lineages in comparison with diploid progenitors, polyploidization has been proposed as a major mechanism of sympatric speciation. Indeed, recent studies show that most (up to 100%) angiosperms have suffered genome duplication across their evolutionary history (Soltis et al. 2003, Wood et al. 2009).

Besides, the mere detection of individuals with different ploidy levels with important applications in areas as biotechnology and plant breeding, in ecology and evolution, the interest has been focused on the dynamics of polyploidization, in particular the origin, establishment and maintenance of polyploids, and how these may evolve into different species (Soltis et al. 2003, Mable 2004). In these areas, polyploids have been approached mainly in three ways: ploidy variation on different spatial and temporal scales; roles of allo- and auto-polyploidization, unreduced gamete production, and single vs multiple origins in the formation of polyploids; and phenotypic and ecological traits of neo- and established polyploids, and the role they play in the ecology and evolution of cytotypes, populations and species (Kron et al. 2007). Using flow cytometry, it has been possible to characterize the geographical distribution of cytotypes (within species or between closely related species), with the main advantage, in comparison with related techniques (*e.g.*, chromosome counting), being the larger number of individuals that can be analysed in a relatively short period of time. As a consequence, the number of studies focused in the study of polyploids has increased significantly in the later years (*e.g.*, Burton and Husband 1999, Bureš et al. 2003, Baack 2004, Pecinka et al. 2006).

Most of the studies so far are focused at population level to describe the patterns of cytotypes distribution. Baack in 2004, observed mixed populations of *Ranunculus adoneus*, with implications for the intercytotype interaction and subsequent

cytotype segregation of this species (Baack et al. 2004); while Lampert and co-authors, found fluctuations in cytotype's frequency over time in *Poecilia formosa* (Lampert et al. 2005).

Still, polyploidy and its origin can be quite complex to entangle, as it envolves the duplication of the genome from the same species (autopolyploidy) or by the combination of genomes from two parental species (allopolyploidy). The first phenomenon is generaly assumed as a mere duplication of the genome and thus, it has less incompatibility and fertility problems, being more easily detected using flow cytometry. In the case of allopolyploids, the process is more complex, and it may bear more fertility problems. Also, its detection using flow cytometry, and the identification of the diploid progenitors (Bennert et al. 2005), usually implies some background knowledge and the analysis of many species from a particular genus. Still, success in not always guaranteed, especially when diploid species have little variation in DNA content or when extinct *taxa* are involved (Kron et al. 2007).

Also, FCM has been been important to address studies focused on evaluating the prevalence of multiple versus single origins of *taxa*, and the role of unreduced gametes in polyploidy formation. Indeed this last process is the most important one in neopolyploids generation (Mable 2004), and therefore it is essential to understand the rates of unreduced gamete production. This can be done using flow cytometry by screening the ploidy level among progeny in controlled crosses. Burton and Husband (2001), in *Chamerion angustifolium*, observed the production of monoploid, diploid, and tetraploid gametes, and a significant role of triploid hybrids in the formation of tetraploids. FCM also enabled to directly measure unreduced microgamete production (still, this is more challenging), and thus to disentangle the rates of gamete production from differential fitness among cytotypes in progeny (Suda et al. 2007b). Some laboratory studies of the effect of experimental treatments on unreduced gamete production suggest an important potential use in natural populations: the comparison of variation in unreduced gamete production to variation in environmental factors (Kron et al. 2007).

In ecological studies, polyploidy is related with phenotypic and fitness differences between polyploids and their diploid progenitors, ecological correlates of ploidy distribution, and the relationship between polyploidy and invasiveness (Thompson and Lumaret 1992). The evidence that the cytotypes differ phenotypically suggest that their differences may have ecological implications (Ramsey and Schemske 2002). Various ecological correlates to polyploidy have been described (Soltis et al. 2003), and

cytotype surveys using FCM frequently include some environmental data (e.g., Baack 2004).

1.3. Flow cytometry: general principles and methods

Flow cytometry is a technique initially developed for rapid counting and analyses of blood cells in clinical research and practice in the late 1950s (Shapero 2007). Later, this technique was applied to a diverse array of scientific areas, including botany, after the discovery of new fluorochromes and development of convenient protocols (Doležel et al. 2007). It is based on the analyses of the optic properties (light dispersion and fluorescence) of particles as they flow individually in a liquid suspension through a light source at high speed (10^2 – 10^3 particles/second), ensuring that particle analysis is completely random, without any subjectivity. This powerful technique enables the analysis of multiple parameters on individual particles in heterogeneous populations. These features make flow cytometry a pivotal technique in areas like diagnostic clinic, with immunophenotyping and cancer research being among the most popular appplications (Doležel 1997).

One vital concept of cytometry is hydrodynamic focusing, a physic phenomenon that near the interrogation point, assures a laminar flow and consequently that particles pass through this point one by one. This is done by an increase in the speed of a sheath fluid, so that it is higher than that of the sample fluid, and thus, both fluids never mix and the flow becames laminar. When the samples, in stream, intercept the laser (interrogation point) a photonic dispersion and/or emission of fluorescence occurs, and this varies directly with the particles characteristics (Côrte-Real et al. 2002). Cytometers can usually detect particles with sizes ranging between 1 and 50 microns in diameter.

The acquired parameters are light scatter and fluorescence. Among light scatter, the light that is dispersed at low angles is named forward light scatter (FS) and is known to be proportional to the particles size. The dispersion of light at wider angles is called side light scatter (SS) and is caused by granularity and the structural complexity of the particles. Fluorescence is the term used to describe the excitation of a fluorophore to a higher energy level followed by the return to the initial energy level, resulting in the emission of light. The energy in the emitted light is dependent on the energy level to which the fluorophore is excited, and that light has a specific wavelength and consequently a specific colour. Any cell component (*e.g.*, DNA, antibodies) binded to a specific fluorescent molecule can be quantified and analysed,

as long as the fluorophore is excited by the light source available in the flow cytometer. As cytometers are usually equipped with a series of filters and mirrors, in most cases, the fluorescence of up to four fluorophores can be analysed at the same time, besides FS and SS.

The light is delivered to appropriate detectors, usually photomultipliers, where the light signal is converted into an analogical signal in the form of a voltage pulse that can be further analysed, after a digital conversion, in a computer (Loureiro 2007). Data is usually obtained in the form of one parameter histograms or two parameters cytograms in specific FCM software. In here quantitative and statistical data are easily obtained for each parameter.

In plants, the main advancements were made after the ingenious method of Galbraith et al. (1983), which develop an easy and quick protocol to isolate and stain plant nuclei. In short, the plant material is chopped in a nuclear isolation buffer (Figure 1). The isolated nuclei are filtered and then stained with a DNA-selective fluorochrome, usually propidium iodide (intercalary fluorochrome) or DAPI (AT-specific fluorochrome). Usually in 10 minutes the sample is ready to be analysed and for most species the resulting fluorescence histograms are of excellent quality. Using this protocol, most plant tissues can be used, but the most used one are the leaves (Loureiro et al. 2010).

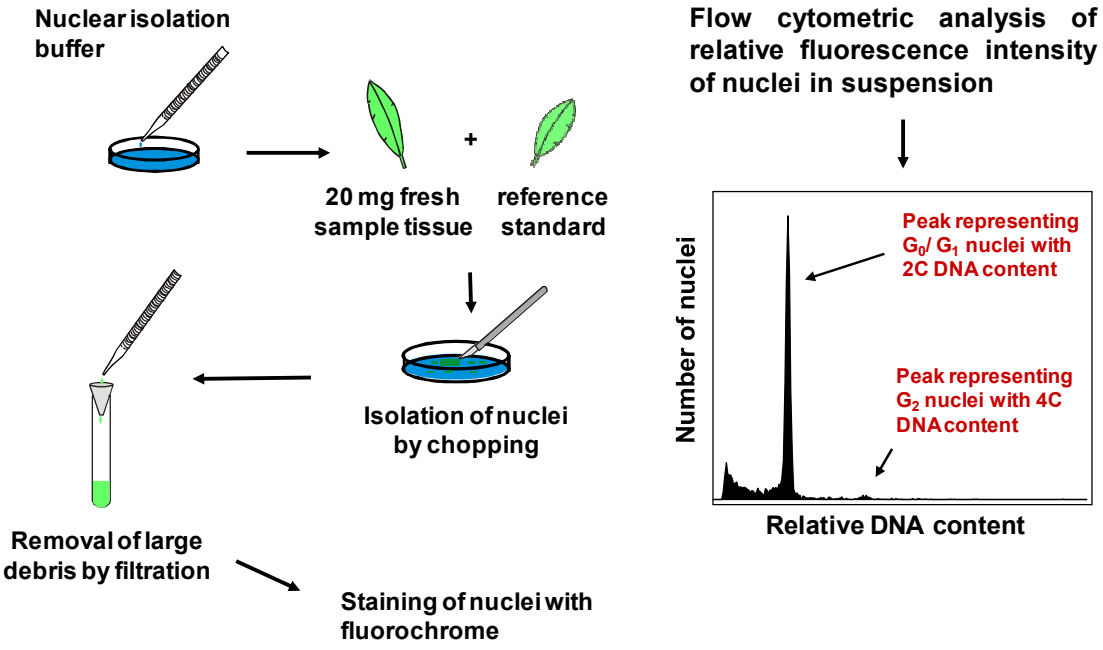

Figure 1. Diagram of the sample preparation procedure for FCM DNA measurements (Galbraith et al. 1983). Figure adapted from the website: http://www.ueb.cas.cz/Olomouc1/LMCC/lmcc.html.

A typical DNA fluorescence histogram comprises a prominent peak corresponding to nuclei in the G_0/G_1 phase of the cell cycle (with a 2C DNA content), a small peak that correspond to nuclei in the G_2 phase (with a 4C DNA content) and some signals in between that correspond to nuclei in the S phase (nuclei synthesising new DNA) (Figure 1). As there are many sources of variation in the FCM analyses, wider peaks and background debris are usually observed (Loureiro 2007), thus the quality of a nuclear suspension is best evaluated by analysing the histogram of relative nuclear DNA content, and the respective CV value of each FL peak. Histograms of good quality should contain a low background debris and symmetrical G_0/G_1 and G_2 peaks with low variation (Doležel and Bartoš 2005).

However, there are also some drawbacks of the use of FCM to analyse plant nuclei. The tissues of several plant species have cytosolic compounds that are released during the nuclear isolation and are known to interfere with fluorescence of nuclear DNA. This artifact may lead to erroneous estimations of nuclear DNA content, and has lead to several controversies. The most well known example is the case of *Helianthus annuus*, where it initially appeared that light exposure could alter the DNA content of this species (Price and Jonhston 1996). However, this intraspecific variation in genome size was not light induced, but because sunflower plants, after light exposure, produce high amounts of cytosolic compounds that interfere with propidium iodide staining, giving a false variation. Other cases followed in the future and this interference is well documented nowadays (Noirot et al. 2002, Loureiro et al. 2006).

Despite of this, the advantages of FCM clearly surpass those of related techniques as chromosome counting and Feulgen microdensitometry (in here the effect of cytosolic compounds is also a problem, Greilhuber 1988): sample preparation is easy, convenient and rapid, only small amounts of plant material are usually required for sample preparation, a variety of plants tissues can be used for sample preparation, it enable the detection of subpopulations and after the equipment is purchased, the demand in consumables is relatively small, which results in relatively low running costs (Loureiro 2007).

Still one should not forget that for most studies it is important to complement FCM results with other cytological techniques as chromosome counting, fluorescence in situ hybridization (FISH) and genomic in situ hybridization (GISH) (Bennett and Leitch 2005).

1.4. Study area

The complex and dynamic paleogeography of the Mediterranean region offers an ideal framework and a natural laboratory to study the diversification patterns of plant lineages. Indeed, Mediterranean flora is characterized by numerous radiations within several plant groups and, thus, it is an important diversification center (e.g., Vargas et al. 2004; Lobo et al. 2001), with Iberian Peninsula in particular being classified as one of the most diverse and biologically richest region of the world (Médail and Quézel, 1997; Myers et al. 2000).

As described above, the ability of flowering plants to give rise to new genetic entities through polyploidization events may have contributed to a rapid taxonomic diversification, eventually being involved on the genesis of the high biodiversity and the numerous endemisms currently observed in some regions of the world. In this context, it might be hypothesised that regions with highly diversified floras, such as the Iberian Peninsula, could be characterized by a high proportion of polyploid species. Considering the paleogeography of this region and the potential broad-scale effects on gene regulation and developmental processes, the Tertiary Iberian orogeny and the subsequent glaciations may have led to important changes in climate and topography (Thompson 1999), which may have provided opportunities for the establishment of polyploids through hybridization and other mechanisms, in the contact zones between existing species. Immediate shifts after polyploidization in morphology, breeding system may have confered reproductive isolation and subsequently lead to diversification (Ramsey and Schemske 1998, Otto and Whitton 2000, Adams and Wendel 2005). On the other hand, polyploids are described as having broader ecological tolerances and thus may have a greater ability to exploit disturbed or novel niches than their diploid progenitors (Levin 2002). This advantage will be most effective when habitats undergo important changes in abiotic conditions or become newly available for colonization (Morton 1993), as has been the case of the Iberian Peninsula.

1.5. Study group

Scrophulariaceae belongs to the order Lamiales (*sensu* Olmstead et al. 1993, Angiosperm Phylogeny Group 1998) which contains several large and well-known families with both tropical and temperate distributions, comprising about 12.3% of the eudicot diversity. The age estimations for Lamiales range from approximately 97 to 74 Myrs ago (Bremer et al. 2004). As traditionally circumscribed (e.g., von Wettstein, 1891), Scrophulariaceae is the largest of Lamiales families and have a worldwide

distribution. The limits of Scrophulariaceae have long been problematic (Thieret 1967, Olmstead 2002) and recent molecular studies using DNA sequences of plastid genes revealed at least five distinct monophyletic groups leading to the desintegration of the traditional classification in, at least, 6 families (Olmstead et al. 2001). Members of the classical Scrophulariaceae are currently found in Scropulariaceae *s.s.*, Plantaginaceae and Orobanchaceae (the latter two contain most of the *taxa* that have moved), as well as in Stilbaceae, Phrymaceae and Linderniaceae (Olmstead et al. 2001, Angiosperm Phylogeny Group II 2003). Despite of this new and currently dynamic state of Scrophulariaceae *s.l.*, most of the tribes *sensu* von Wettstein (1891) that continue to be currently accepted have been shown to be monophyletic through molecular studies (e.g., tribe *Anthirrhineae*, Vargas et al. 2004) and can thus be considered natural groups.

Scrophulariaceae *s.l.* comprise plants that can be perennial, annual or biannual herbs, rarely shrubs. Their stem has circular or square section; the leaves are highly diverse and can be simple to pinnately dissected, alternate or opposite, entire to deeply indented, rarely in verticils and basal rosettes, with petiole or sessile, without stipules. The flowers are usually tubular with bilateral symmetry (zygomorphic), rarely actinomorphic, arranged in inflorescences or solitary; inflorescences can be spikes or panicles with bracts; the calyx is usually penta-lobated, more rarely bi-, tetra- or octo-lobated; the corolla is simpatelous with (4)5(8) petals forming a tube more or less developed, sometimes gibbous or ending in a spur; androecium is composed by 2-8 stamens, more frequently 4 didynamous stamens, and in some species staminoids are present; pollen grains are tricolporate or tricolpate; the gynoecium is bicarpelar and the ovary is superior and bilocular, with axial placentation and numerous ovules. Finally, the fruits are dehiscent capsules and the seeds have endosperm (Benedí et al. 2009).

A huge diversity in corolla shape, colour and functioning can be observed in the family. Characteristically, Scrophulariaceae individuals have gullet blossoms specially adapted to bee pollinations. The features that characterize these flowers are zygomorphy, corolla tube shaped, frequently with a landing platform (lower lip), nectar produced at the base of the corolla, sexual organs placed under the upper lip in some cases closed corolla dependent on bee weight to trigger the opening of the corolla (Faegri and van der Pijl 1979).

In Iberian Peninsula, Scrophulariaceae *s.l.* are represented by 323 species distributed in 33 genera (Benedí et al. 2009). In Portugal, the diversity is lower, with 116 species distributed in 27 genera being found. Most species are ruderal and can be easily found in disturbed lands; still, there are several species listed in the red lists, and

thus in need of special protection (*e.g.*, *Anarrhinum longipedicellatum*, Habitats Directive ofNATURA2000; VV AA, 2000).

Considering the oldness of this family, that there is almost no available information on genome size for any *taxa* of this family (but see Albach and Greilhuber 2004), that there are several records in the literature pointing to the possible existence of polyploids within and between species of Scrophulariaceae (e.g., *Antirrhinum*, *Digitalis* and *Veronica*) and that, in case polyploids are found, many *taxa* present large attractive flowers, ideal for reproductive isolation studies, it seemed that this family had all the necessary attributes for a large-scale cytogenetic-based survey.

1.6. Objectives

The main objectives of the present Master Thesis were to assess the value of genome size as a taxonomic marker and the role of polyploidy as a process of genesis and maintenance of plant diversity in Scrophulariaceae *s.l.* in Iberian Peninsula. To achieve these goals we proposed three levels of fundamental research: 1) assessment of chromosomes numbers, genome size and polyploidy incidence in Scrophulariaceae *taxa* from Iberian Peninsula by performing a review of the bibliographic literature; 2) estimation of genome size in a diverse array of *taxa* from several key genera; and 3) assessment of cytotype diversity through large-scale screenings in natural populations. The results will increase the basic scientific knowledge on genome evolution and polyploid incidence in the Scrophulariaceae from Iberian Flora, providing important background information for subsequent studies, namely taxonomic studies in some interesting groups and focused on the ecological significance of genome size and polyploidy and their importance in plant diversification in this region.

Chapter 2

MATERIALS and METHODS

2.1 Plant material

Plant samples from 59 *taxa* of the Scrophulariaceae *s.l.* family were collected from several field locations in Portugal and Spain (Figure 2 and Appendix 1). Seeds from some *taxa* were kindly provided by *index semina* of several Iberian research institutions (Appendix 1). Also, seeds from the reference standards were kindly provided by the Laboratory of Molecular Cytogenetics and Cytometry, Institute of Experimental Botany, Olomouc, Czech Republic.

Field collections were carried out in 2009 and 2010 during the flowering season (March to August) of the studied *taxa*. For that, when possible, natural populations presenting at least 30 individuals were selected. In each population, leaves and/or seeds from up to 30 individuals (one leaf/seed per individual) were collected and stored in a hermetic plastic bag identified with an ID code referring to the collector, population and *taxa*. Samples were kept at 4ºC in a refrigerator until analysis. Photographs were made of each population and *taxa*, with special care being taken on photographing the habit and flowers of each taxon. Also, GPS coordinates were acquired and registered in a field book, being given the same ID code of the plastic bag. Voucher specimens were also collected and kept in the Herbarium of the University of Coimbra (COI). In the laboratory, each taxon was identified following Benedí et al. (2009).

An extensive bibliographic review on chromosome counts, localities and genome size of the studied species was carried out. For chromosome information and localities the following bibliography or online databases were used: Flora Iberica (Benedí et al. 2009), Tropicos® (http://www.tropicos.org/), Anthos (http://www.anthos.es/), BioDiversity4all (http://www.biodiversity4all.org/; for localities only) and M. Queirós printed files database available at the Department of Life Sciences, Faculty of Sciences and Technology, University of Coimbra. For genome size information the Kew Plant DNA C-values Database (release 5.0, December 2010; http://data.kew.org/cvalues/) was used.

Seeds from Scrophulariaceae *taxa* and from reference standards were sown in plastic cuvettes filled with commercial peat. Plastic cuvettes were put in a greenhouse operating at 20 ± 2 ºC and with a photoperiod of 16h/8h (light/dark) and a light intensity of 530 ± 2 μmol m^{-2} s^{-1}.

2.2. Calibration of reference standards for genome size estimations

According to Greilhuber et al. (2007), before starting a large scale survey of genome size variation across a whole family, due to need of using multiple reference standards, one should consider to select a primary reference standard (e.g., *Pisum sativum*) and calibrate the genome size of the remaining standards against it. Considering the nature of the present study it was decided to follow this recommendation. In brief, *Pisum sativum* cv. Ctirad (2C = 8.76 pg DNA, Doležel et al. 1998) was selected as a primary reference standard and all the other standards, with exception of *Raphanus sativus* cv. Saxa, were calibrated against it (see details below for the flow cytometric procedure). In the case of *R. sativus*, to avoid possible problems of non-linearity due to dissimilar genome sizes between this species and *Pisum sativum*, *Solanum lycopersicon* cv. Stupické was used. For each species pair (*P. sativum* / other standard), five replicates were done in three different days. When the coefficient of variation (CV) of the five genome size estimates was larger than 2% the most dissimilar values were discarded and new replicates were made until this threshold quality value was achieved.

2.3 Genome size and ploidy level estimations using flow cytometry

For flow cytometric (FCM) analyses of genome size and ploidy level, leaves from field collected or seed germinated plants were used as plant material. In one taxon (*Rhinanthus minor*), because all plants were in fructification and no leaves were available, seeds were used as an alternative tissue. Nuclear suspensions were prepared according to Galbraith et al. (1983), by chopping approximately 50 mg of plant material of the study species and 50 mg of leaves of the internal reference standard with a sharp razor blade in a glass Petri dish containing 1 mL of WPB buffer (0.2 M Tris.HCl, 4 mM $MgCl_2.6H_2O$, 1 % Triton X-100, 2 mM EDTA $Na_2.2H_2O$, 86 mM NaCl, 10 mM metabissulfite, 1% PVP-10, pH adjusted to 7.5 and stored at 4ºC; Loureiro et al. 2007b). For each taxon/population, after the first sample, if necessary the chopping intensity and amount of plant material was adjusted in order to have a rate of 20-50 nuclei/s in the subsequent replicates. In samples with a large amount of cytosolic compounds, the chopping intensity was reduced to avoid their release from the cells and, thus prevent or minimize their negative effect on nuclear fluorescence (Loureiro et al. 2006). When a novel species was analysed for the first time, a quick overview of the literature was made using the FLOWER database (http://flower.web.ua.pt/) to determine which reference standard was previously used. In case no information was available, *P. sativum* or *S. lycopersicon* were chosen as

starting standards, and according to the obtained results the most appropriate standard for each taxon was selected. When possible and justifiable, the same standard was used for all the *taxa* of each genus. Nuclear suspensions were then filtered through an 50 μm nylon filter and 50 μg/mL of propidium iodide (PI, Fluka, Buchs, Switzerland) and 50 μg/mL of RNAse (Fluka, Buchs, Switzerland) were added to sample tubes to stain the DNA and avoid staining of double stranded RNA, respectively. Samples were kept at room temperature and analysed within a 10 minute period in a Partec CyFlow Space flow cytometer (Partec GmbH, Görlitz, Germany) equipped with a 532 nm green solid-state laser, operating at 30 mW. Integral fluorescence and fluorescence height and width emitted from nuclei were collected through a 620 nm band-pass interference filter. For a given taxon the amplifier system was set to a constant voltage and gain, throughout the whole estimates. Each day, prior to analysis, the instrument stability and linearity was checked either with fluorescent beads or using PI stained nuclei isolated from *P. sativum*. The analyses were only started when CV values were below 3%. If this was not achieved both a cleaning procedure and an adjust of the position of the flow chamber with respect to the incident laser were made until the desired CV values were obtained.

The results were acquired using the Partec FloMax software (v. 2.5) in the form of five graphics: fluorescence pulse integral in linear scale (FL); forward light scatter (FS) *vs.* side light scatter (SS), both in logarithmic (log) scale; FL *vs.* time; FL *vs.* fluorescence pulse height; FL *vs.* SS in log scale. In most samples, in the latter graphic, a polygonal region was defined to include only intact nuclei. This region was used as a gating in all the other graphics. This procedure enabled to remove partial nuclei, nuclei with associated cytoplasm and other debris from analysis, and thus obtain a clearer view on the position of each peak. In FL histograms, linear regions were delimited around the G_0/G_1 peaks of both sample and standard nuclei. This enabled to obtain information on the number of nuclei, mean fluorescence and CV value of the fluorescence of the particles included in this region. At least 1,300 nuclei in each G_0/G_1 peak were analysed per sample. For each taxon, only genome size estimates presenting a CV value below 5 % were considered. Samples with higher CV values were discarded and a new sample was prepared. For some *taxa* with high amounts of cytosolic compounds it was not possible to achieve such CV values, and thus a higher CV threshold was considered acceptable (8 %).

For each population, the genome size of 3 individuals was analysed. The ploidy level of the remaining individuals was analysed by pooling 5-6 individuals in one sample, *i.e.*, a piece of leaf of each individual (the same quantity of each leaf) was

added to the Petri dish and chopped together with the reference standard. In case different ploidy levels were detected (given by multiple G_0/G_1 peaks), the number of nuclei present in each G_0/G_1 peak enabled to determine the number of individuals with each ploidy level.

Ploidy level analyses consisted on determining the ratio between the mean FL of sample's nuclei and the mean FL of standard's nuclei. The holoploid genome size in pg (2C; complete genome size, *sensu* Greilhuber et al. 2005) of each individual was estimated according with the following formula:

$$\text{2C nuclear DNA content} = \frac{\text{Scrophulariaceae sp. } G_0/G_1 \text{ peak mean FL}}{\text{reference standard } G_0/G_1 \text{ peak mean FL}} \times \text{nuclear DNA content of reference standard.}$$

The monoploid genome size (1Cx; *i.e.*, the single genome with *x* chromosomes, of which there are two per unreplicated nucleus in a diploid individual and several in a polyploid individual, *sensu* Greilhuber et al. 2005) of all species was also calculated by dividing the holoploid genome size (2C) by the supposed ploidy level of each *taxa*, both in mass values (pg) and Mbp (1 pg = 978 Mbp; Doležel et al. 2003).

2.4 Statistical analysis

Descriptive statistics of genome size were calculated for each species (mean, standard deviation of the mean and coefficient of variation of the mean). For genera with more than one species, box plots with mean and standard deviation of the mean were computed using Microsoft Excel 2007. 1Cx values of *Veronica* species were also plotted.

Differences in genome size among families considering the newly established boundaries (*i.e.*, Scrophulariacea *s.s.*, Orobanchaceae and Plantaginaceae) were assessed using a non-parametric Kruskal-Wallis one-way ANOVA on ranks (normality and homoscedasticity was not achieved even after data transformation). For genus with more than one sampled species (*Anarrhinum*, *Antirrhinum*, *Digitalis*, *Linaria*, *Misospates*, *Pedicularis*, *Scrophularia*, *Verbascum* and *Veronica*) differences in nuclear DNA content within and between species were evaluated using Statistica v.8.0. A similar analysis was performed between closely related genera (*Bartsia*, *Nothobartsia* and *Parentucellia,* and *Odontite vs Odontitella*). For variables that were normally distributed and homoscedastic a t-test (comparisons between two groups) or a one-way ANOVA (comparisons of more than two groups) were followed. In *Linaria* spp. and *Veronica* spp., data transformations ($\log_{10}$ and square root, respectively) had to be

used to achieve normality and homoscedasticity. In *Scrophularia* sp., due to failure in achieving homoscedasticity (even after data transformation), a non-parametric Kruskal-Wallis one-way ANOVA on ranks was used. When statistically significant differences were detected, either a multiple comparison Tukey-Kramer test (for parametric data) or a Dunn's method (for non-parametric data) were applied to determine which groups presented significantly different values.

In *Veronica* spp., a linear regression analysis and a Pearson correlation were performed between mean nuclear DNA content and chromosome numbers of each *taxa*.

Chapter 3
RESULTS

The bibliographic review on chromosome counts and localities of the 116 Scrophulariaceae *s.l. taxa* present in Portugal is presented in Appendix 2. From the analysis of this data, 28 *taxa* presented more than one value of chromosome counts, despite that only in 10 *taxa* (8.6% of the total) this may represent different ploidy levels (*e.g.*, *Digitalis purpurea* subsp. *purpurea* and *Odontites vernus*, both with 2*x* and 4*x*; *Veronica cymbalaria* and *V. hederifolia*, both with 2*x* and 3*x*). For the remaining *taxa*, usually differences of 2 or more chromosomes are reported, but never an additional full set of chromosomes (Appendix 2). With respect to the distribution, 56% of the *taxa* occurred in more than 3 Portuguese provinces; still, 25 *taxa* were restricted to one province. Field observations revealed that, as reported in the literature, most species are ruderal and can be easily found in disturbed lands, in many cases due to anthropogenic activities (e.g., *Digitalis purpurea* subsp. *purpurea*, *Verbascum* sp.).

The available data in the literature on genome size variation in Scrophulariaceae (minimum value of 0.34 pg/2C in *Torenia baillonii*, Kikuchi et al. 2006, and maximum value of 14.30 pg/2C in *Collinia verna*, Greenlee et al. 1984) suggested the need to use multiple standards to obtain reliable estimations of nuclear DNA content in this family. Therefore, a calibration of the necessary reference standards against *Pisum sativum* (the primary reference standard) was performed. The results of this calibration are given in Table 2. Due to a slightly lower value assumed to *P. sativum*, as recommended by Greilhuber et al. (2006), all the calibrated 2C values were lower than the original ones provided by the Laboratory of Molecular Cytogenetics and Cytometry (Olomouc, Czech Republic).

Table 2. DNA reference standards used in the present study, with both the original and calibrated 2C DNA values

Species	Cultivar	Calibrated 2C DNA content (pg)	Original 2C DNA content (pg)[1]	Reference
Pisum sativum	'Ctirad'	8.76*	9.09	Doležel et al. 1998
Zea mays	'CE-777'	5.30	5.43	Lysák and Doležel 1998
Glycine max	'Polanka'	2.39	2.50	Doležel et al. 1994
Solanum lycopersicum	'Stupické'	1.92	1.96	Doležel et al. 1992
Raphanus sativus	'Saxa'	1.08	1.11	Doležel et al. 1992

[1] Nuclear DNA content established using human male leukocytes (2C = 7.0 pg DNA; Tiersch et al. (1989)) as a primary reference standard.

[*] *Pisum sativum* cv. Ctirad (Doležel et al. 1998) served as a primary reference standard, with a 2C value of 8.76 pg as recommended by Greilhuber et al. (2006).

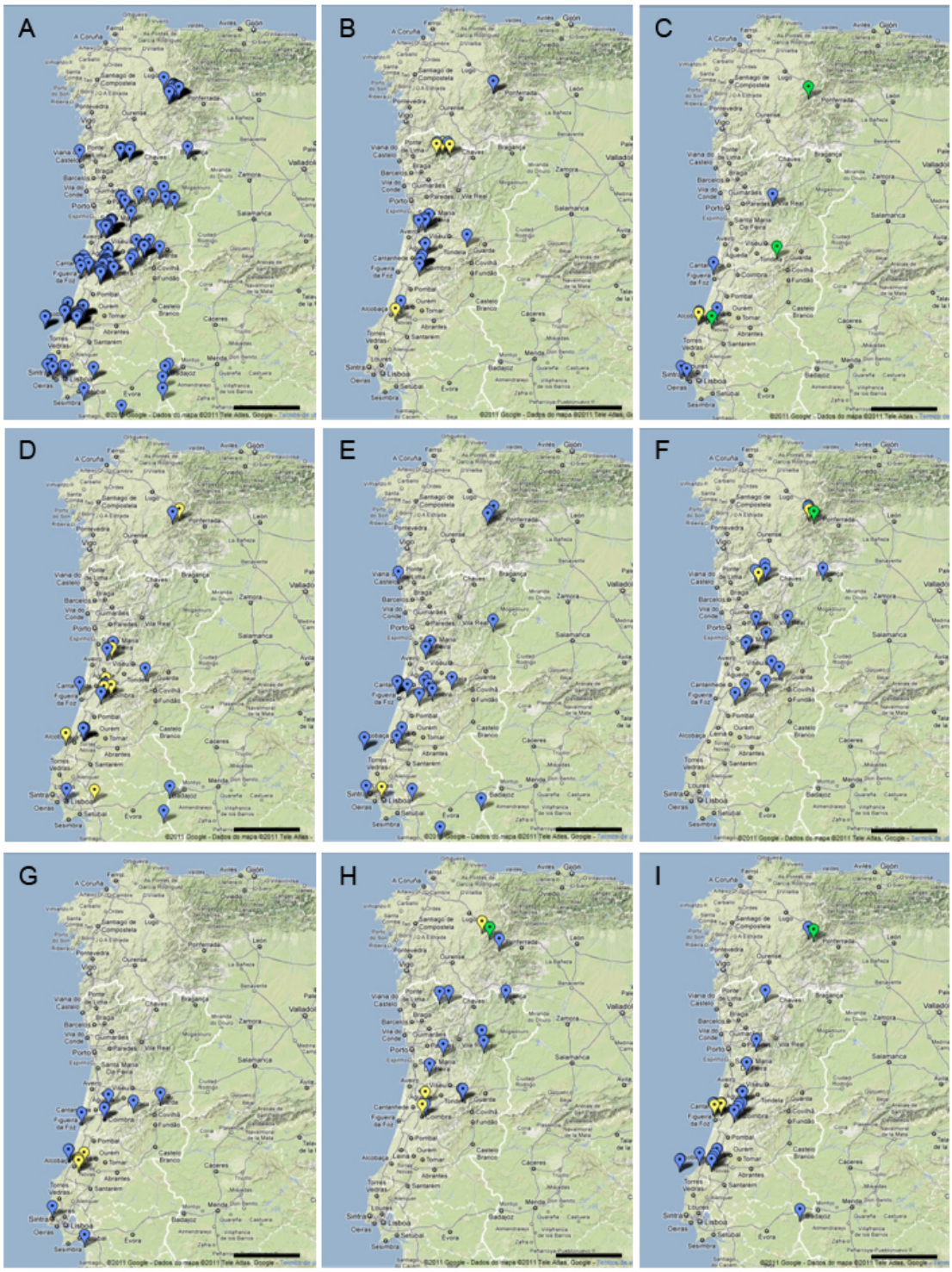

Figure 2. Maps with the location of the sampled populations: (**A**) all collected populations; (**B**) populations of the genera *Pedicularis* (yellow) and *Veronica* (blue); (**C**) populations of the genera *Nothobartsia* (yellow), *Bartsia* (blue) and *Parentucellia* (green); (**D**) populations of the genera *Misopates* (yellow) and *Verbascum* (blue); (**E**) populations of the genera *Kickxia* (yellow) and *Linaria* (blue); (**F**) populations of the genera *Melampyrum* (yellow), *Anarrhinum* (blue) and *Euphrasia* (green); (**G**) populations of the genera *Chaenorhinum* (yellow) and *Antirrhinum* (blue); (**H**) populations of the genera *Cymbalaria* (yellow), *Digitalis* (blue) and *Rhinanthus* (green); (**I**) populations of the genera *Odontite* (yellow), *Scrophularia* (blue) and *Odontitella* (green). The black bar represents 100 km.

The use of FCM enabled to perform a large scale analysis of 17 genera of Scrophulariaceae, comprising 59 species (51% of the species known to occur in Portugal) and a total of 162 populations (Table 3 and Figure 2). From the 59 sampled species, 86% are first estimations of genome size (Table 3). With a few exceptions (e.g., *Veronica micrantha*) the overall quality of the results, as given by the CV values of G_0/G_1 peaks and by the background debris, was good, with mean CV values below 5 % being achieved in most *taxa* (Table 3 and Figure 3).

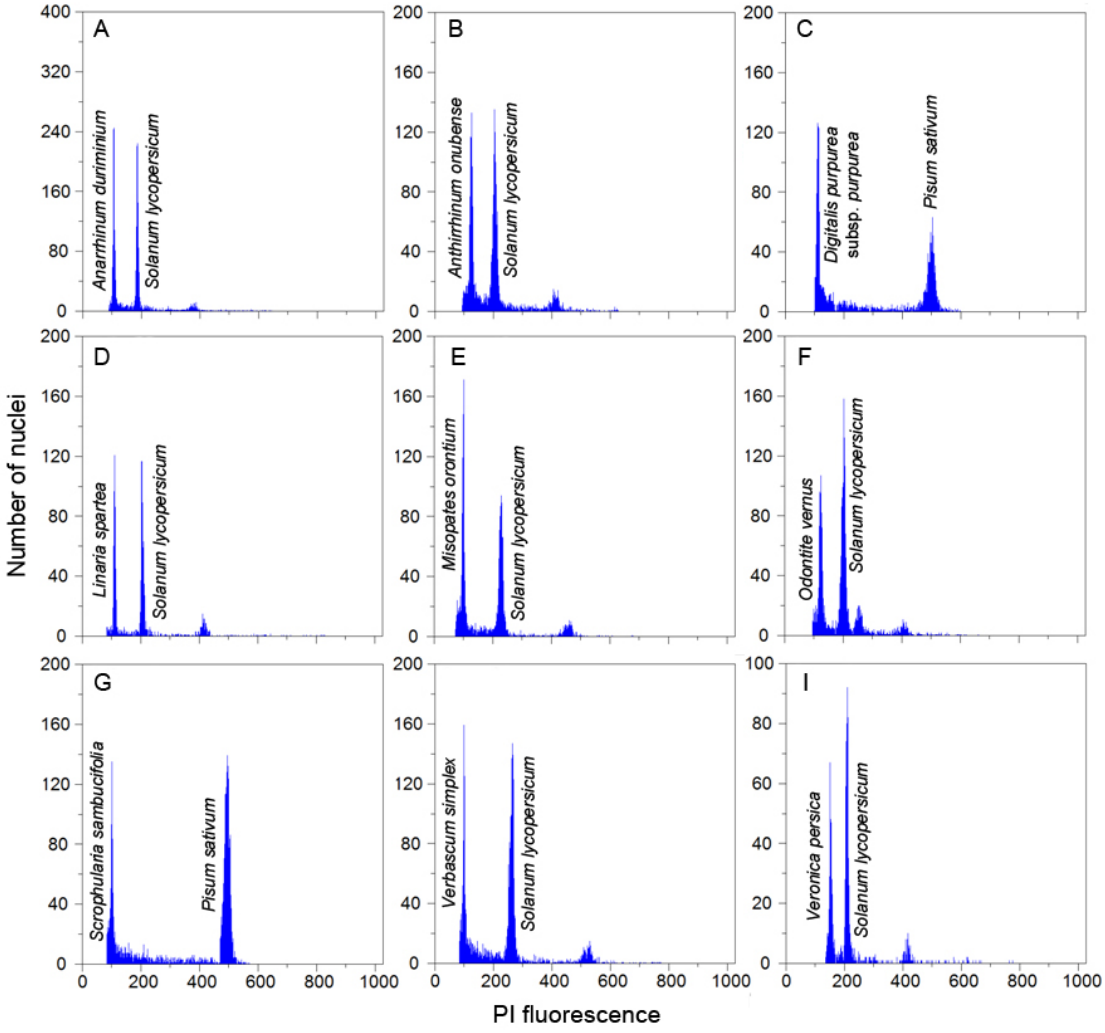

Figure 3. Flow cytometric histograms of relative PI fluorescence intensity obtained after simultaneous analysis of nuclei isolated from the internal reference standard and from the Scrophulariaceae species: (**A**) G_0/G_1 peaks of *Anarrhinum longipedicellatum* and *Solanum lycopersicum*; (**B**) G_0/G_1 peaks of *Antirrhinum onubense* and *S. lycopersicum*; (**C**) G_0/G_1 peaks of *Digitalis purpurea* subsp. *purpurea* and *Pisum sativum*; (**D**) G_0/G_1 peaks *of Linaria spartea* and *S. lycopersicum*; (**E**) G_0/G_1 peaks of *Misopates orontium* and *S. lycopersicum*; (**F**) G_0/G_1 peaks of *Odontite vernus* and *S. lycopersicum*; (**G**) G_0/G_1 peaks of *Scrophularia sambucifolia* and *P. sativum*; (**H**) G_0/G_1 peaks of *Verbascum simplex* and *S. lycopersicum*; (**I**) G_0/G_1 peaks *Veronica persica* and *S. lycopersicum*. In histograms A, B, D-F, H and I it is possible to observed the G_2 peak of the internal reference standard; additionally, in F it is also possible to observe the G_2 peak of *O. vernus* (third peak). Also, please note the overall good quality of the histograms, as defined by the narrow G_0/G_1 peaks and by the low amount of background debris.

Table 3. Nuclear DNA content estimations in the studied *taxa* of Scrophulariaceae *s.l.*

Taxon	Family *s.s.*	Genome size (2C, pg)		Genome size (1Cx, Mbp)[1]	Genome size (1Cx, pg)	FL CV (%)	Ploidy level	Standard[2]	n G.s.	n total	n Pop.	Origin	Previous estimations
		Mean ± SD	CV (%)										
Anarrhinum bellidifolium	Plantaginaceae	1.13 ± 0.03	3.1	553	0.56	3.68	2n=2*x*	S	32	157	11	POP+IS	First estimation
Anarrhinum duriminium	Plantaginaceae	1.11 ± 0.02	2.2	545	0.56	5.61	2n=2*x*	S	12	35	3	POP+IS	First estimation
Anarrhinum longipedicelatum	Plantaginaceae	1.12 ± 0.02	1.4	547	0.57	4.05	2n=2*x*	S	7	40	3	POP	First estimation
Antirrhinum cirrhigerum	Plantaginaceae	1.21 ± 0.01	0.8	594	0.59	5.95	2n=2*x*	S	3	30	1	IS	First estimation
Antirrhinum graniticum	Plantaginaceae	1.18 ± 0.05	3.9	576	0.59	5.06	2n=2*x*	S	3	30	1	IS	First estimation
Antirrhinum linkianum	Plantaginaceae	1.23 ± 0.03	2.7	600	0.61	4.66	2n=2*x*	S	17	66	7	POP+IS	First estimation
Antirrhinum meonanthum	Plantaginaceae	1.20	-	588	0.61	5.36	2n=2*x*	S	1	1	1	IS	First estimation
Antirrhinum onubense	Plantaginaceae	1.18 ± 0.01	1.1	579	0.61	4.15	2n=2*x*	S	3	20	1	POP	First estimation
Bartsia trixago	Orobanchaceae	1.85 ± 0.08	4.1	907	0.93	3.91	2n=2*x*	G/S	17	103	6	POP+IS	First estimation
Chaenorhinum origanifolium	Plantaginaceae	1.13 ± 0.02	1.3	555	0.57	3.47	2n=2*x*	S	9	61	3	POP	First estimation
Cymbalaria muralis subsp. *muralis*	Plantaginaceae	0.99 ± 0.02	2.5	482	0.49	5.12	2n=2*x*	S	9	48	3	POP	First estimation
Digitalis mariana subsp. *heywoodii*	Plantaginaceae	1.12	-	546	0.56	6.77	2n=2*x*	S	1	1	1	IS	First estimation
Digitalis purpurea subsp. *purpurea*	Plantaginaceae	1.87 ± 0.05	2.6	917	0.94	3.64	2n=2*x*	B/P	19	168	11	POP+IS	2C = 2.45 pg[A]
Digitalis thapsi	Plantaginaceae	2.08	-	1017	1.04	5.90	2n=2*x*	Z	1	1	1	IS	First estimation
Euphrasia minimus	Orobanchaceae	1.29 ± 0.02	1.3	631	0.65	3.02	2n=2*x*	S	3	30	1	POP	First estimation
Kickxia spuria subsp. *integrifolia*	Plantaginaceae	1.64 ± 0.02	1.2	801	0.82	3.44	2n=2*x*	S	4	17	1	IS	First estimation
Linaria aeruginea subsp. *aeruginea*	Plantaginaceae	1.29 ± 0.01	0.9	629	0.64	3.83	2n=2*x*	S	2	5	1	POP	First estimation
Linaria amethystea subsp. *amethystea*	Plantaginaceae	1.05 ± 0.01	0.6	514	0.53	3.67	2n=2*x*	S	3	30	1	POP	First estimation
Linaria diffusa	Plantaginaceae	1.15 ± 0.00	0.4	560	0.57	2.97	2n=2*x*	S	2	15	1	POP	First estimation

Linaria incarnata	Plantaginaceae	1.13 ± 0.00	0.3	552	0.56	3.31	2n=2*x*	S	2	15	1	POP	First estimation
Linaria polygalifolia subsp. *polygalifolia*	Plantaginaceae	1.32 ± 0.04	2.7	647	0.66	4.17	2n=2*x*	S	12	70	4	POP	First estimation
Linaria saxatilis	Plantaginaceae	1.21	-	583	0.60	6.12	2n=2*x*	S	1	1	1	IS	First estimation
Linaria spartea	Plantaginaceae	1.11 ± 0.05	4.1	541	0.55	4.05	2n=2*x*	S	29	149	9	POP+IS	First estimation
Linaria supina	Plantaginaceae	1.30 ± 0.03	2.7	637	0.65	3.76	2n=2*x*	S	9	57	3	POP	First estimation
Linaria thriornithophora	Plantaginaceae	2.66 ± 0.08	3.0	1302	1.33	3.03	2n=2*x*	S	14	98	5	POP	First estimation
Melampyrum pratense subsp. *latifolium*	Orobanchaceae	15.69 ± 0.19	1.2	7670	7.84	3.27	2n=2*x*	P	6	46	2	POP	First estimation
Misopates calycinum	Plantaginaceae	0.88 ± 0.04	4.4	431	0.44	4.04	2n=2*x*	S	3	26	1	POP	First estimation
Misopates orontium	Plantaginaceae	0.88 ± 0.04	4.3	431	0.44	4.91	2n=2*x*	S	19	97	7	POP	First estimation
Northobartsia asperrima	Orobanchaceae	1.55 ± 0.02	1.3	756	0.77	3.58	2n=2*x*	S	3	27	1	POP	First estimation
Odontite vernus	Orobanchaceae	1.16 ± 0.02	1.8	569	0.58	4.05	2n=2*x*	S	3	30	1	POP	First estimation
Odontitella virgata	Orobanchaceae	4.27 ± 0.02	0.5	2088	2.13	2.93	2n=2*x*	G/S	6	60	2	POP	First estimation
Parentucellia viscosa	Orobanchaceae	2.72 ± 0.06	2.0	1331	1.36	2.83	2n=2*x*	S	6	24	3	POP	First estimation
Pedicularis sylvatica subsp. *lusitanica*	Orobanchaceae	5.95 ± 0.15	2.5	2909	2.97	2.42	2n=2*x*	S	8	29	3	POP	First estimation
Pedicularis sylvatica subsp. *sylvatica*	Orobanchaceae	5.61 ± 0.02	0.3	2744	2.81	3.15	2n=2*x*	S	2	2	1	POP	First estimation
Rhinanthus minor	Orobanchaceae	2.81 ± 0.08	2.8	1373	1.40	5.26	2n=2*x*	P	3	20	1	POP	2C = 7.9 pg[B]
Scrophularia auriculata subsp. *auriculata*	Scrophulariaceae	1.79 ± 0.01	1.6	877	0.90	3.98	2n=2*x*	P	8	8	1	IS	First estimation
Scrophularia frutenscens	Scrophulariaceae	1.34 ± 0.03	2.5	653	0.67	5.55	2n=2*x*	P	7	34	3	POP+IS	First estimation
Scrophularia grandiflora	Scrophulariaceae	1.94 ± 0.07	6.6	948	0.97	4.13	2n=2*x*	B/G/P	19	51	6	POP+IS	First estimation
Scrophularia herminii	Scrophulariaceae	2.56 ± 0.07	2.7	1252	1.28	6.15	2n=2*x*	P	3	16	1	IS	First estimation
Scrophularia lyrata	Scrophulariaceae	3.19 ± 0.05	0.7	1561	1.60	4.54	2n=2*x*	P	3	15	1	POP	First estimation
Scrophularia nodosa	Scrophulariaceae	1.19 ± 0.01	0.6	583	0.60	6.71	2n=2*x*	P	2	2	1	IS	First estimation
Scrophularia sambucifolia subsp. *sambucifolia*	Scrophulariaceae	1.86 ± 0.04	2.0	909	0.93	4.33	2n=2*x*	P	5	5	1	POP	First estimation
Scrophularia scorodonia	Scrophulariaceae	2.11 ± 0.05	2.2	1034	1.06	4.57	2n=2*x*	B/G/P	19	106	7	POP+IS	First estimation

Scrophularia sublyrata	Scrophulariaceae	2.22 ± 0.12	5.5	1087	1.11	5.95	2n=2*x*	B	5	15	2	POP	First estimation
Verbascum levanticum	Scrophulariaceae	0.75 ± 0.02	2.9	368	0.38	5.57	2n=2*x*	R	3	16	1	POP	First estimation
Verbascum litigiosum	Scrophulariaceae	0.76 ± 0.03	4.2	370	0.38	3.48	2n=2*x*	S	3	30	1	POP	First estimation
Verbascum pulverulentum	Scrophulariaceae	0.78 ± 0.02	2.2	383	0.39	4.15	2n=2*x*	S	3	30	1	POP	First estimation
Verbascum simplex	Scrophulariaceae	0.74 ± 0.02	2.8	361	0.37	3.70	2n=2*x*	S	12	70	4	POP	First estimation
Verbascum sinuatum	Scrophulariaceae	0.77 ± 0.04	4.7	379	0.39	5.08	2n=2*x*	S	17	121	6	POP+IS	First estimation
Verbascum virgatum	Scrophulariaceae	1.44 ± 0.02	1.5	350	0.36	3.51	**2n=4*x***	S	4	11	2	POP+IS	First estimation
Veronica acinifolia	Plantaginaceae	1.24 ± 0.01	0.7	608	0.62	3.73	2n=2*x*	S	3	3	1	POP	First estimation
Veronica arvensis	Plantaginaceae	0.91 ± 0.01	1.6	447	0.46	3.92	2n=2*x*	S/R	9	58	3	POP	2C = 0.66 pg[C]
Veronica chamaedrys* subsp. *chamaedrys	Plantaginaceae	3.72 ± 0.02	0.6	607	0.62	3.70	**2n=6*x***	G/S	3	30	1	POP	2C = 2.98 pg[C]
Veronica hederifolia	Plantaginaceae	4.16 ± 0.08	2.0	678	0.69	2.84	**2n=6*x***	B	3	8	1	POP	2C = 2.82 pg[C]
Veronica micrantha	Plantaginaceae	2.15 ± 0.04	1.7	525	0.54	7.56	**2n=4*x***	P	3	17	1	IS	First estimation
Veronica officinalis	Plantaginaceae	2.10 ± 0.06	2.9	514	0.53	3.98	**2n=4*x***	B/P	12	51	4	POP	First estimation
Veronica peregrina* subsp. *peregrina	Plantaginaceae	1.96 ± 0.06	2.9	479	0.49	4.02	2n=4*x*	B	3	8	1	POP	2C = 1.90 pg[C]
Veronica persica	Plantaginaceae	1.40 ± 0.03	2.4	342	0.35	4.70	2n=4*x*	S	24	105	7	POP	2C = 1.55 pg[C]
Veronica polita	Plantaginaceae	0.77 ± 0.01	1.5	378	0.39	4.61	2n=2*x*	S	6	18	2	POP	2C = 0.84 pg[C]

The values are given as mean and standard deviation of the mean of the holoploid genome size (2C, pg) of individuals of each species. For every species, the monoploid genome size (1Cx) in Mbp and in mass values (pg), the mean coefficient variation (CV, %) of G_0/G_1 peaks, the supposed ploidy level, the reference standard used to estimate the genome size (standard), the number of individuals analysed for genome size (n G.s.), the total number of analysed individuals (n total), the total number of analysed populations (n Pop.) and the origin of plant material (POP, natural populations; IS, index seminum) are also given. Also, for each species, previous genome size estimations and original references are provided ([A] Mowforth 1986, [B] Nagl and Fusenig 1979, [C] Albach and Greilhuber 2004). In bold, the new DNA ploidy levels are highlighted.

[1] 1 pg = 978 Mbp (Doležel et al. 2003)

[2] R, *Raphanus sativus* cv. Saxa; S, *Solanum lycopersicum* cv. Stupické; G, *Glycine max* cv. Polanka; B, *Bellis perenis*; Z, *Zea mays* cv. CE-777; P, *Pisum sativum* cv. Ctirad; see Table 2 for genome size estimations of the standards).

Among the sampled species, a genome size variation of 21.6-fold was found, with the lowest mean value being obtained for *Verbascum simplex* (2C = 0.74 ± 0.02 pg) and the highest one for *Melampyrum pratense* subsp. *latifolium* (2C = 15.69 ± 0.19 pg). Still, according with the genome size categories defined by Leitch et al. (1998), 89.8 % of the *taxa* have a very small genome (2C ≤ 2.8 pg), 8.5 % have a small genome (2.8 pg < 2C ≤ 7.0 pg) and 1.7 % have an intermediate genome (7.0 pg < 2C ≤ 28.0 pg) (Figure 4). No species with large (28.0 pg < 2C ≤ 70.0 pg) and very large (2C > 70.0 pg) genome sizes were detected. No significant differences in genome size were obtained among families considering the newly established taxonomy (H_2 = 5.47, P = 0.065).

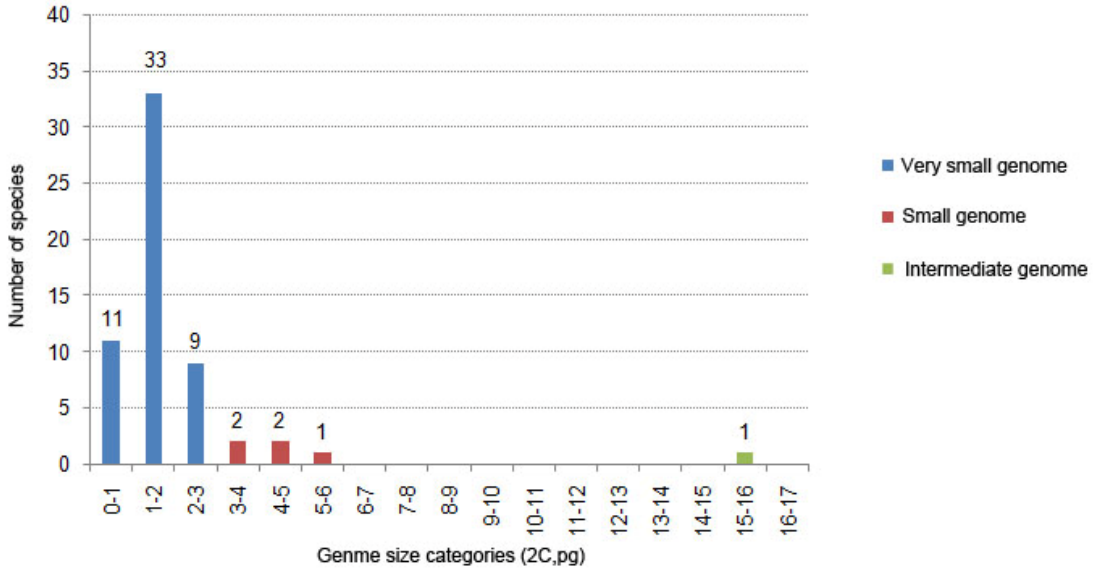

Figure 4. Distribution of genome size according with genome categories (1 pg difference). Colors represent the categories defined by Leitch et al. (1998).

A detailed analysis of the variation of genome size within each genus, revealed that no statistically significant differences were detected in genome size in *Anarrhinum* (3 analysed species; F_2 = 1.51, P = 0.230), *Anthirrhinum* (5 analysed species; F_4 = 2.39, P = 0.082) and *Misopates* (2 analysed species; t = 0.01, P = 0.991) (Figure 5 and Appendix 3). In all the other genera (*Digitalis*, *Linaria*, *Pedicularis*, *Scrophularia*, *Verbascum* and *Veronica*) statistically significant differences were observed (see details of the tests in Appendix 3, Figures 6-9), with genome size being an important character to separate at least two *taxa* within each genus.

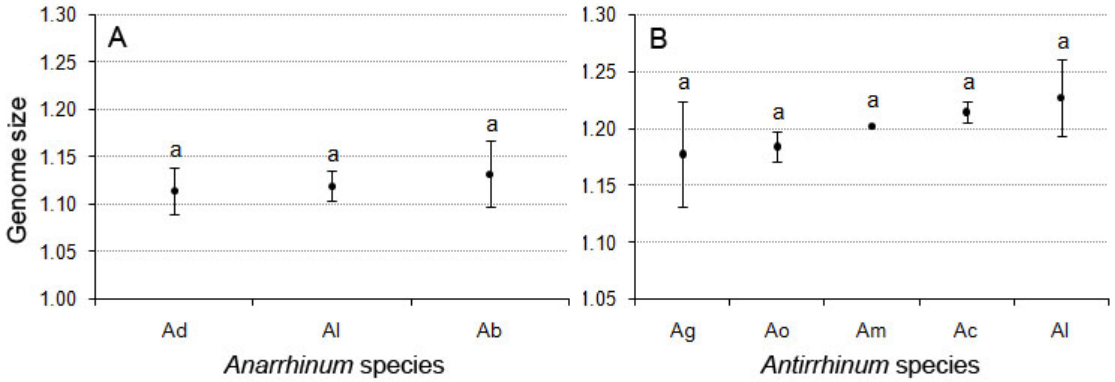

Figure 5. Genome size variation (mean and standard deviation of the mean) in: (**A**) *Anarrhinum* sp. (Ad, *Anarrhinum duriminium*; Al, *Anarrhinum longipedicelatum*; Ab, *Anarrhinum bellidifolium*) and (**B**) *Antirrhinum* sp. (Ag, *Antirrhinum graniticum*; Ao, *Antirrhinum onubense*; Am, *Antirrhinum meonanthum*; Ac, *Antirrhinum cirrhigerum*; Al, *Antirrhinum linkianum*). Groups followed by the same letter are not significantly different at $P < 0.05$.

In *Scrophularia*, *Verbascum* and *Veronica* genera genome size differences are due to different numbers of chromosomes among *taxa*. While, in *Verbascum* and *Veronica* the species with the highest number of chromosomes present the largest genome, in *Scrophularia*, the species with the higher number of chromosomes (*S. auriculata* subsp. *auriculata*, 78-88 chromosomes, Appendix 2), presented a lower genome size (2C = 1.79 pg) than expected, if a positive linear correlation between chromosome numbers and genome size would be considered (data not shown). Also, despite many *Scrophularia* species were not statistically different, due to dissimilar and non-overlapping values of genome size, it was possible to use this character to separate several *taxa* (e.g., *S. nodosa*, *S. frutescens*, *S. hederifolia*, S. *lyrata*; $P < 0.05$; Figure 6C). In *Verbascum*, despite it was possible to statistically distinguish *V. virgatum* from all the other analysed *taxa*, the latter ones had very similar genome sizes being statistically similar ($P < 0.05$; Figure 6D).

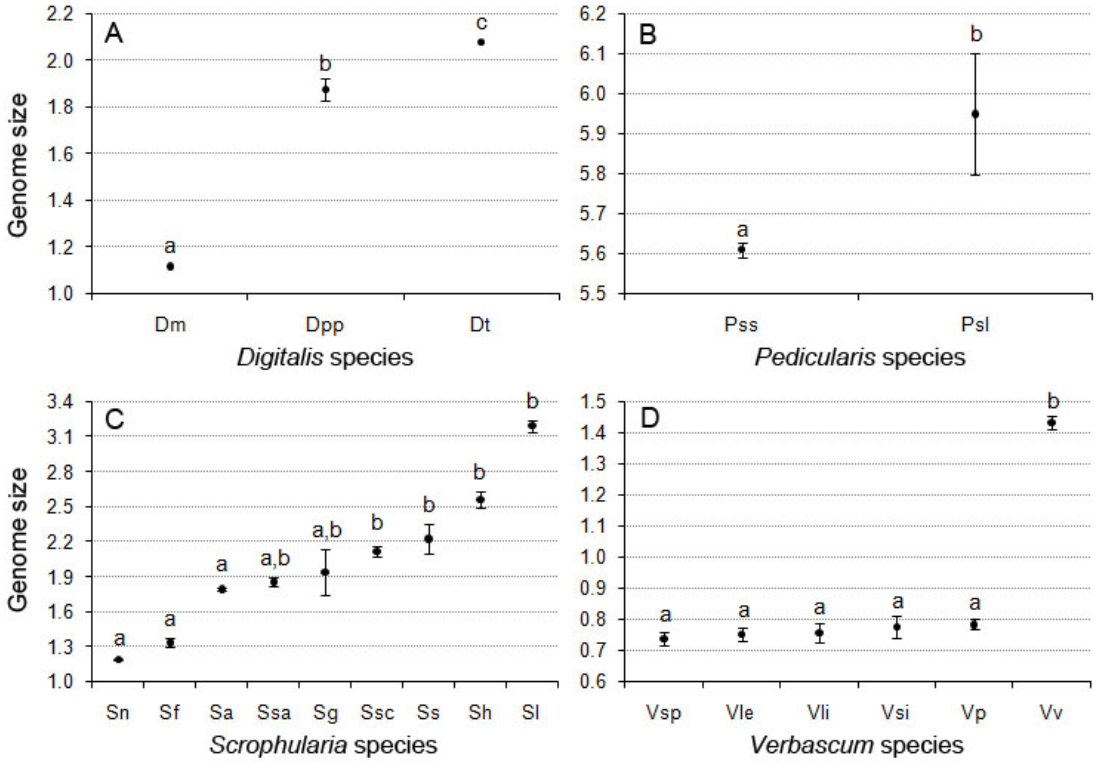

Figure 6. Genome size variation (mean and standard deviation of the mean) in Scrophulariaceae genera: (**A**) *Digitalis* sp. (Dm, *D. mariana*; Dpp, *D. purpurea* subsp. *purpurea*, Dt, *D. thapsi*); (**B**) *Pedicularis* sp. (Pss, *P. sylvatica* subsp. *sylvatica*; Psl, *P. sylvatica* subsp. *lusitanica*); (**C**) *Scrophularia* sp. (Sn, *S. nodosa*; Sf, *S. frutescens*; Sa, *S. auriculata* subsp. *auriculata*; Ssa,*S. sambucifolia* subsp. *sambucifolia*; Sg, *S. grandiflora*; Ssc, *S. scorodonia*; Ss, *S. sublyrata*; Sh, *Scrophularia herminii*; Sl, *S. lyrata*); (**D**) *Verbascum* sp. (Vsp, *V. simplex*; Vle, *V. levanticum*; Vli, *V. litigiosum*; Vsi, *V. sinuatum*; Vp, *V. pulverulentum*; Vv, *V. virgatum*). Different letters represent groups that are significantly different ($P < 0.05$).

In *Veronica*, with the exception of *V. officinalis* and *V. micrantha*, all the other analysed *taxa* were significantly different in genome size ($P < 0.05$; Figure 7A). With respect to the monoploid genome size (1Cx), some variation was also observed (ranging from 0.33 pg in *V. peregrina* subsp. *peregrina* to 0.69 pg in *V. hederifolia*) (Figure 7B). A detailed analysis considering the sections to which the analysed *taxa* belong revealed more homogenous 1Cx values in section *Veronica* than in section *Pocilla* (Figure 7B). The linear regression analysis between chromosome numbers and genome size, revealed a positive correlation between these characters, with a relatively high R^2 value of 0.7229 (Figure 7C). A Pearson correlation analysis confirmed this result (correlation coefficient of 0.85, $P < 0.05$).

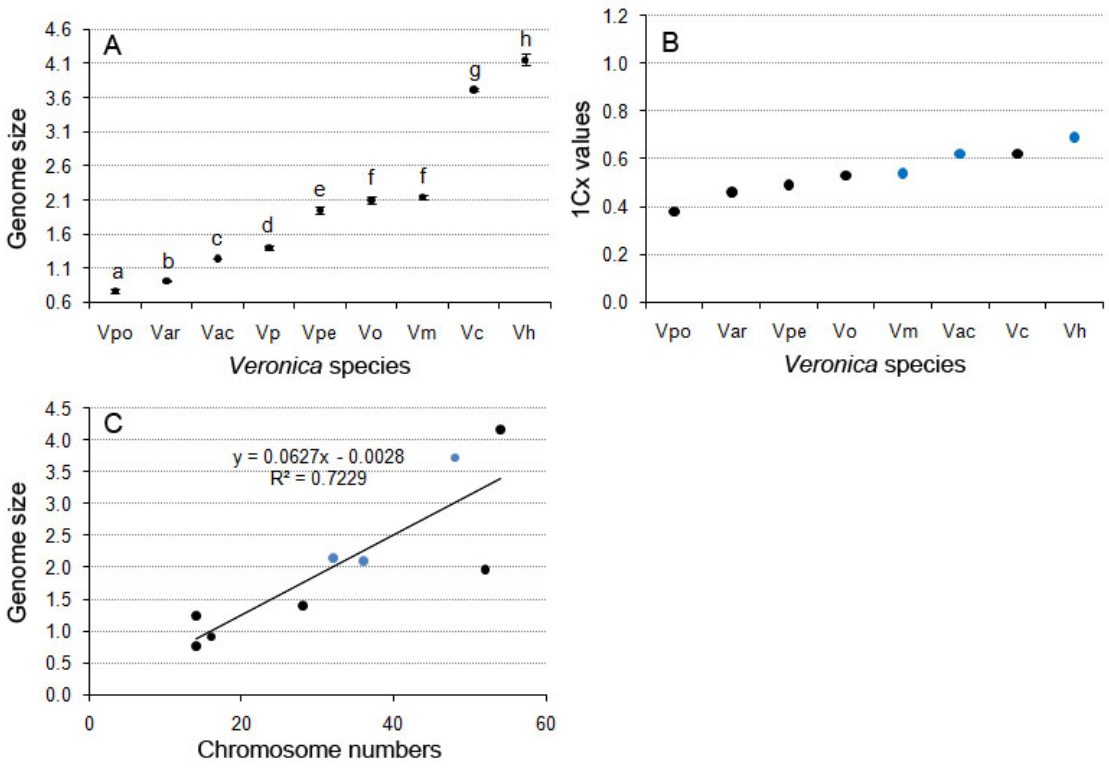

Figure 7. Genome size variation in *Veronica* sp.: (**A**) Genome size (mean and standard deviation of the mean); (**B**) 1Cx values; (**C**) linear regression analyses between mean nuclear DNA content and chromosome number (the linear regression equation and the R^2 coefficient are also provided). Species: Vpo, *V. polita*; Var, *V. arvensis*; Vac, *V. acinifolia*; Vp, *V. persica*; Vpe, *V. peregrina* subsp. *peregrina*; Vo, *V. officinalis*; Vm, *V. micrantha*; Vc, *V. chamaedrys* subsp.*chamaedrys*; Vh, *V. hederifolia*. Different letters represent groups that are significantly different ($P < 0.05$). Black dots represent the species belonging to section *Pocilla* and the blue dots represent species that belong to section *Veronica*.

In the case of *Digitalis*, *Linaria* and *Pedicularis*, according with the literature, all the analyzed *taxa* within each genus present the same number of chromosomes (56, 14 and 16, respectively, Appendix 2). However, regardless of the same number of chromosomes, statistically significant differences in genome size were detected (Appendix 3), being possible to separate the three analysed *taxa* of *Digitalis* ($P < 0.05$; Figure 6A) and the two analysed *taxa* of *Pedicularis* (Figure 6B), using this character only. In *Linaria*, *L. triornithophora* presented a statistically distinguishable higher genome size than the remaining species ($P < 0.05$; 2C = 2.66 pg; Figure 8A); still, the other species presented dissimilar but close values of genome size ranging from 1.05 pg/2C in *L. amethystea* subsp. *amethystea* to 1.32 pg in *L. polygalifolia* subsp. *polygalifolia*, not all distinguishable statistically ($P < 0.05$; Figure 8B).

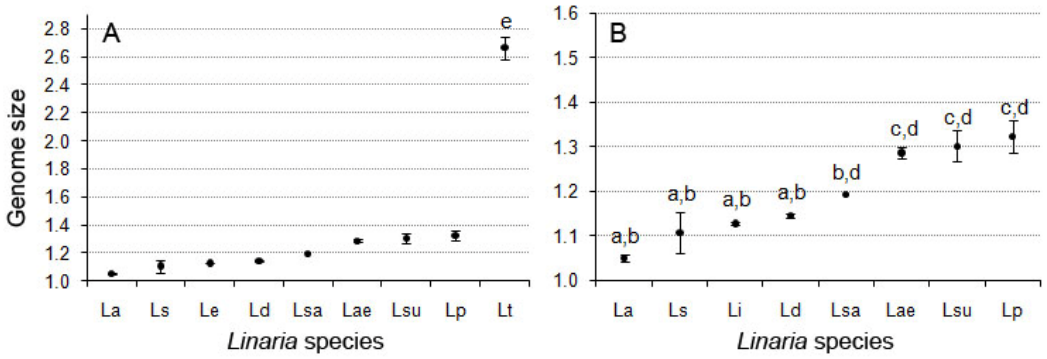

Figure 8. Genome size variation (mean and standard deviation of the mean) in *Linaria* genus: (**A**) La, *Linaria amethystea* subsp. *amethystea*; Ls, *Linaria spartea*; Li, *Linaria incarnata*; Ld, *Linaria diffusa*; Lsa, *Linaria saxatilis*; Lae, *Linaria aeruginea* subsp. *aeruginea*; Lsu, *Linaria supina*; Lp, *Linaria polygalifolia* subsp. *polygalifolia*; Lt, *Linaria triornithophora*; (**B**) Detail of the figure 9A, excluding *Linaria triornithophora* from the analysis. Different letters represent groups that are significantly different ($P < 0.05$).

A comparison of the genome size values obtained for the species of the closely related genera *Bartsia*, *Nothobartsia* and *Parentucellia*, revealed that the analysed *taxa* have statistically significant differences ($F_2 = 300.76$, $P < 0.001$; Appendix 3), with *N. asperrima* presenting the lowest genome size value (2C = 1.55 pg) and *P. viscosa* presenting the highest genome size with 2C = 2.72 pg ($P < 0.05$; Figure 9A). A similar analysis but with species from the closely related genera *Odontites* and *Odontitella*, also revealed statistically significant differences between the two genera (t = 206.23, P < 0.001; Figure 9B; Appendix 3).

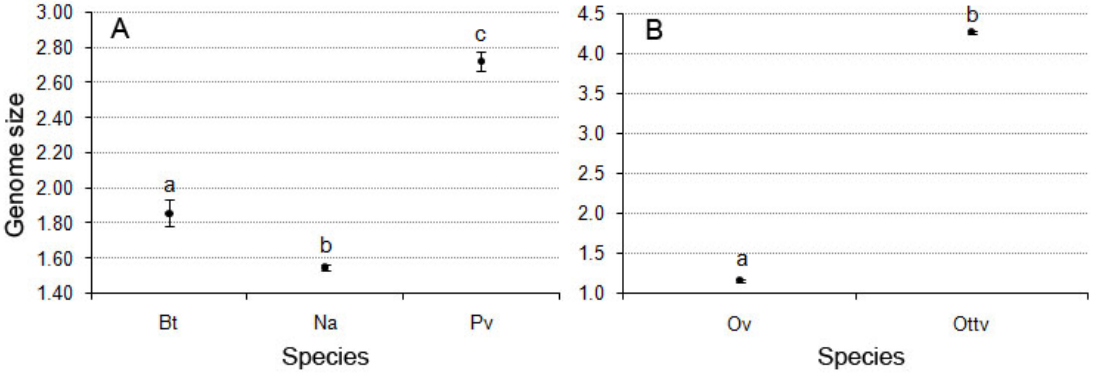

Figure 9. Genome size variation (mean and standard deviation of the mean) in the following Scrophulariaceae species: (**A**) *Bartsia trixago* (Bt), *Nothobartsia asperrima* (Na) and *Parentucellia viscosa (*Pv); (**B**) *Odontite vernus* (Ov), *Odontitella virgata* (Ottv). Different letters represent groups that are significantly different ($P < 0.05$).

To evaluate the incidence of intraspecific variation of genome size, *i.e.*, variation among populations of the same *taxa*, a thorough analysis was made for the genera where statistically significant differences among *taxa* were obtained and where more populations were collected (*Digitalis*, *Linaria*, *Scrophularia, Verbascum and Veronica*; Tables 4-8, respectively). In *Digitalis*, the main focus was on *Digitalis purpurea* subsp. *purpurea* where a CV value of 2.6 % and a DNA range of 1.76 to 2.06 pg/2C (n = 10 populations) were obtained (Table 4). Still, in all populations except population 9 (MC92) the CV values were low (< 2.0 %). In population 9 a CV value of 5.6 % was obtained, reflecting three very dissimilar genome size estimates within this population. These results at population level reflect the scenario obtained for this *taxon*, with 3 main groups of estimates differing of approximately 0.11 pg being obtained between populations (population 10 with mean value of 1.79 pg, populations 5-8 with mean values of 1.90 pg and populations 1-4 with mean values of 2.01 pg; Table 4).

In *Linaria*, usually low CV values of genome size (< 3.0 %) were obtained among populations of the same *taxa*. Still, in *L. triornithophora* and *L. spartea* higher CV values were found. In the case of *L. triornithophora*, this is mainly due to the estimates obtained for two of the individuals of population 3 (MC71, Table 5), which had approximately 8% less DNA than the mean value obtained for all the other individuals. With respect to *L. spartea*, a variation on DNA values of 14% was found, with individuals from population 5 (MC44, Table 5) presenting the lowest values, and those from populations 1 and 6 (SC14 and MC88, respectively) presenting the highest values. Still, these were among the populations with the higher homogeneity in genome size among individuals (Table 5).

In *Scrophularia*, some heterogeneity in genome size values was found within some species (e.g., *S. grandiflora*), but mostly among the estimates within populations (Table 6). For example, it is curious to notice that despite the mean value for all the populations of *S. frutescens* and *S. sublyrata* was very similar, in population 1 of *S. frutescens* and in both populations of *S. sublyrata*, highly dissimilar values of genome size were obtained, with CV values higher than 5 % being obtained. In *S. grandiflora*, both differences among populations (e.g., population 5 with mean 2C of 2.01 pg *vs.* population 6 with mean 2C of 1.87 pg) and within populations (e.g., population 4 with 2C values spanning from 1.80 to 1.94 pg) wre obtained. Also, in *S. scorodonia*, in one population (MC51) the obtained estimates of genome size are among the most extreme found for this species (Table 6).

In *Verbascum*, some heterogeneity in genome size estimates was observed, with seven out 15 populations of different *taxa* presenting genome size CV values higher

than 3.5%. This was mostly due to dissimilar estimations within populations (e.g., population 3 of *V. sinuatum* with a DNA range from 0.74 to 0.81 pg/2C, and the only population of V. litigiosum with a DNA range between 0.72 and 0.78 pg/2C; Table 6) and not among populations (Table 7). In *Veronica*, all the species and populations presented CV values below 3.5 %, revealing a good homogeneity in the genome size values both between and within populations (Table 8).

Concerning the incidence of polyploidy in Scrophulariaceae, contrarily to what was expected, at least for some *taxa* (Appendix 2), no different cytotypes were detected among any of the 162 surveyed populations in any of the 59 *taxa*. Still, as referred above, within some genera (e.g., *Veronica*, *Verbascum*) there are species with different DNA ploidy levels (*sensu* Suda et al. 2006). In the particular case of Veronica, according with the 1Cx analysis presented in Figure 7C, three novel DNA ploidy levels were assumed, namely hexaploid populations in *V. chamaedrys* subsp. *chamaedrys* and *V. hederifolia*, and tetraploid populations in *V. officinalis and V. micrantha* (Table 3).

Table 4. Nuclear DNA content estimations in the studied *taxa* of *Digitalis*.

ID code	*Taxon*	Pop.	Genome size (2C, pg)				Genome size (1Cx, Mbp)[1]	Genome size (1Cx, pg)	FL CV (%)	Ploidy level
			Mean ± SD	CV (%)	Min.	Max.				
COI8	*Digitalis mariana* subsp. *heywoodii*	1	1.19	-	-	-	581	0.594	6.77	2n=2*x*
SC29	*Digitalis purpurea* subsp. *purpurea*	1	2.00 ± 0.03	1.3	1.98	2.03	976	0.998	2.51	2n=2*x*
SC34	*Digitalis purpurea* subsp. *purpurea*	2	2.03 ± 0.04	1.8	1.99	2.06	993	1.015	2.57	2n=2*x*
SC39	*Digitalis purpurea* subsp. *purpurea*	3	1.97 ± 0.01	0.5	1.96	1.98	962	0.983	2.71	2n=2*x*
SC41	*Digitalis purpurea* subsp. *purpurea*	4	2.02 ± 0.02	1.2	1.99	2.03	988	1.010	3.03	2n=2*x*
UTP20	*Digitalis purpurea* subsp. *purpurea*	5	1.90 ± 0.02	1.2	1.88	1.93	931	0.952	6.94	2n=2*x*
MC17	*Digitalis purpurea* subsp. *purpurea*	6	1.90 ± 0.01	0.5	1.89	1.91	929	0.950	3.21	2n=2*x*
MC20	*Digitalis purpurea* subsp. *purpurea*	7	1.89 ± 0.01	0.6	1.87	1.89	922	0.943	3.07	2n=2*x*
MC69	*Digitalis purpurea* subsp. *purpurea*	8	1.89 ± 0.03	1.8	1.86	1.93	925	0.946	4.19	2n=2*x*
MC92	*Digitalis purpurea* subsp. *purpurea*	9	1.90 ± 0.11	5.6	1.81	2.02	930	0.950	3.44	2n=2*x*
MC94	*Digitalis purpurea* subsp. *purpurea*	10	1.79 ± 0.03	1.5	1.76	1.81	873	0.893	4.22	2n=2*x*
COI9	*Digitalis thapsi*	1	2.08	-	-	-	1017	1.040	5.90	2n=2*x*

The values are given as mean and standard deviation of the mean of the holoploid genome size (2C, pg) of individuals of each species. For each collection, the ID code and population number (see Appendix 1 for details), the DNA range (Min., minimum genome size; Max., maximum genome size), the monoploid genome size (1C*x*) in Mbp and in mass values (pg), the mean coefficient of variation (CV, %) of G_0/G_1 peaks and the supposed ploidy level are also given.

[1] 1 pg = 978 Mbp (Doležel et al. 2003)

Table 5. Nuclear DNA content estimations in the studied *taxa* of *Linaria*.

ID code	*Taxon*	Pop.	Genome size (2C, pg)				Genome size (1Cx, Mbp)[1]	Genome size (1Cx, pg)	FL CV (%)	Ploidy level	Section, Subsection
			Mean ± SD	CV (%)	Min.	Max.					
MC12	*Linaria triornithophora*	1	2.73 ± 0.02	0.8	2.70	2.75	1333	1.363	4.41	2n=2*x*	*Pelisserianae*
MC19	*Linaria triornithophora*	2	2.71 ± 0.01	0.2	2.70	2.71	1324	1.354	2.71	2n=2*x*	*Pelisserianae*
MC71	*Linaria triornithophora*	3	2.54 ± 0.09	3.6	2.48	2.65	1244	1.272	3.05	2n=2*x*	*Pelisserianae*
SC30	*Linaria triornithophora*	4	2.69 ± 0.01	0.2	2.68	2.69	1315	1.344	2.69	2n=2*x*	*Pelisserianae*
MC99	*Linaria triornithophora*	5	2.64 ± 0.00	0.1	2.64	2.65	1293	1.322	2.33	2n=2*x*	*Pelisserianae*
COI16	*Linaria saxatilis*	1	1.19	-	-	-	583	0.597	6.12	2n=2*x*	*Supinae, Saxatile*
SC28	*Linaria aeruginea* subsp. *aeruginea*	1	1.28 ± 0.01	0.7	1.27	1.29	627	0.641	3.97	2n=2*x*	*Supinae, Supinae*
MC49	*Linaria amethystea* subsp. *amethystea*	1	1.05 ± 0.01	0.6	1.05	1.06	514	0.525	3.67	2n=2*x*	*Supinae, Supinae*
MC14	*Linaria diffusa*	1	1.15 ± 0.00	0.4	1.14	1.15	560	0.573	2.97	2n=2*x*	*Supinae, Supinae*
SC01	*Linaria polygalifolia* subsp. *polygalifolia*	1	1.35 ± 0.01	0.1	1.34	1.36	661	0.676	3.86	2n=2*x*	*Supinae, Supinae*
SC25	*Linaria polygalifolia* subsp. *polygalifolia*	2	1.35 ± 0.01	0.4	1.34	1.36	660	0.674	4.47	2n=2*x*	*Supinae, Supinae*
MC84	*Linaria polygalifolia* subsp. *polygalifolia*	3	1.32 ± 0.01	0.4	1.32	1.33	647	0.662	2.97	2n=2*x*	*Supinae, Supinae*
COI14	*Linaria polygalifolia* subsp. *polygalifolia*	4	1.27 ± 0.00	0.7	1.27	1.27	620	0.634	5.40	2n=2*x*	*Supinae, Supinae*
MC29	*Linaria supina*	1	1.32 ± 0.01	0.8	1.31	1.33	646	0.661	4.20	2n=2*x*	*Supinae, Supinae*
MC63	*Linaria supina*	2	1.26 ± 0.02	1.6	1.24	1.28	616	0.629	3.63	2n=2*x*	*Supinae, Supinae*
MC101	*Linaria supina*	3	1.33 ± 0.01	0.8	1.31	1.33	551	0.564	3.47	2n=2*x*	*Supinae, Supinae*
MC13	*Linaria incarnata*	1	1.13 ± 0.00	0.3	1.12	1.13	552	0.564	3.31	2n=2*x*	*Versicolores, Versicolores*
SC14	*Linaria spartea*	1	1.16 ± 0.01	0.7	1.15	1.17	566	0.579	4.05	2n=2*x*	*Versicolores, Versicolores*
JC03	*Linaria spartea*	2	1.13 ± 0.04	3.1	1.10	1.17	551	0.563	4.04	2n=2*x*	*Versicolores, Versicolores*
JC07	*Linaria spartea*	3	1.13 ± 0.01	1.2	1.13	1.15	555	0.567	4.31	2n=2*x*	*Versicolores, Versicolores*
JC18	*Linaria spartea*	4	1.11 ± 0.02	1.5	1.10	1.13	545	0.557	5.62	2n=2*x*	*Versicolores, Versicolores*
MC44	*Linaria spartea*	5	1.02 ± 0.02	1.8	1.01	1.04	501	0.512	4.75	2n=2*x*	*Versicolores, Versicolores*
MC88	*Linaria spartea*	6	1.16 ± 0.01	0.7	1.15	1.16	566	0.578	4.44	2n=2*x*	*Versicolores, Versicolores*
MC89	*Linaria spartea*	7	1.13 ± 0.02	1.3	1.11	1.14	551	0.563	2.98	2n=2*x*	*Versicolores, Versicolores*
MNHN30	*Linaria spartea*	8	1.07 ± 0.04	3.5	1.03	1.11	525	0.537	2.88	2n=2*x*	*Versicolores, Versicolores*
MNHN32	*Linaria spartea*	9	1.09 ± 0.02	1.7	1.09	1.12	532	0.544	3.08	2n=2*x*	*Versicolores, Versicolores*

The values are given as mean and standard deviation of the mean of the holoploid genome size (2C. pg) of individuals of each species. For each collection, the ID code and population number (see Appendix 1 for details), the DNA range (Min.. minimum genome size; Max.. maximum genome size), the monoploid genome size (1C*x*) in Mbp and in mass values (pg), the mean coefficient of variation (CV. %) of G_0/G_1 peaks and the supposed ploidy level are also given. The section and subsection of each *taxon* is also provided. [1] 1 pg = 978 Mbp (Doležel et al., 2003).

Table 6. Nuclear DNA content estimations in the studied *taxa* of *Scrophularia*.

ID code	*Taxon*	Pop.	Genome size (2C, pg)				Genome size (1Cx, Mbp)[1]	Genome size (1Cx, pg)	FL CV (%)	Ploidy level	Section, Subsection
			Mean ± SD	CV (%)	Min.	Max.					
COI24	*Scrophularia frutescens*	1	1.35 ±0.08	5.6	1.29	1.40	658	0.673	6.64	2n=2*x*	*Caninae, Caninae*
MC81	*Scrophularia frutescens*	2	1.33 ± 0.01	0.9	1.32	1.34	652	0.667	5.85	2n=2*x*	*Caninae, Caninae*
MC86	*Scrophularia frutescens*	3	1.33 ± 0.01	1.4	1.32	1.34	651	0.666	3.57	2n=2*x*	*Caninae, Caninae*
MNHN34	*Scrophularia auriculata* subsp. *auriculata*	1	1.79 ± 0.01	0.7	1.77	1.81	877	0.900	4.54	2n=2*x*	*Scrophularia, Scrophularia*
SC20	*Scrophularia grandiflora*	1	2.00 ± 0.01	0.7	1.99	2.01	977	0.999	3.20	2n=2*x*	*Scrophularia, Scrophularia*
MC18	*Scrophularia grandiflora*	2	1.96 ± 0.02	1.2	1.94	1.99	958	0.980	3.25	2n=2*x*	*Scrophularia, Scrophularia*
MC26	*Scrophularia grandiflora*	3	1.94 ± 0.01	0.7	1.93	1.95	948	0.970	3.49	2n=2*x*	*Scrophularia, Scrophularia*
MC42	*Scrophularia grandiflora*	4	1.88 ± 0.07	3.9	1.80	1.94	921	0.942	3.64	2n=2*x*	*Scrophularia, Scrophularia*
MC50	*Scrophularia grandiflora*	5	2.01 ± 0.04	2.1	1.98	2.06	982	1.004	5.00	2n=2*x*	*Scrophularia, Scrophularia*
UTP22	*Scrophularia grandiflora*	6	1.87 ± 0.05	2.8	1.83	1.94	914	0.935	4.03	2n=2*x*	*Scrophularia, Scrophularia*
UTP23	*Scrophularia herminii*	1	2.56 ± 0.07	2.7	2.50	2.63	1252	1.280	7.49	2n=2*x*	*Scrophularia, Scrophularia*
MC93	*Scrophularia lyrata*	1	3.19 ± 0.05	1.6	3.15	3.25	1561	1.597	3.98	2n=2*x*	*Scrophularia, Scrophularia*
ISAUTL7	*Scrophularia nodosa*	1	1.19 ± 0.01	0.6	1.19	1.20	583	0.298	6.71	2n=2*x*	*Scrophularia, Scrophularia*
MC58	*Scrophularia sambucifolia* subsp. *sambucifolia*	1	1.86 ± 0.04	2.0	1.81	1.91	909	0,929	4.33	2n=2*x*	*Scrophularia, Scrophularia*
MC22	*Scrophularia scorodonia*	1	2.16 ± 0.04	1.9	2.13	2.19	1055	1.079	3.87	2n=2*x*	*Scrophularia, Scrophularia*
MC46	*Scrophularia scorodonia*	2	2.10 ± 0.02	0.9	2.08	2.12	1027	1.050	2.95	2n=2*x*	*Scrophularia, Scrophularia*
MC51	*Scrophularia scorodonia*	3	2.15 ± 0.09	4.1	2.07	2.24	1052	1.076	3.92	2n=2*x*	*Scrophularia, Scrophularia*
MC76	*Scrophularia scorodonia*	4	2.13 ± 0.03	1.3	2.11	2.16	1040	1.063	5.84	2n=2*x*	*Scrophularia, Scrophularia*
MC77	*Scrophularia scorodonia*	5	2.09 ± 0.02	0.9	2.07	2.10	1021	1.044	6.04	2n=2*x*	*Scrophularia, Scrophularia*
MC110	*Scrophularia scorodonia*	6	2.09 ± 0.04	2.1	2.04	2.13	551	0.564	5.46	2n=2*x*	*Scrophularia, Scrophularia*
UPT24	*Scrophularia scorodonia*	7	2.09 ± 0.02	0.8	2.07	2.10	1022	1.045	2.57	2n=2*x*	*Scrophularia, Scrophularia*
JC02	*Scrophularia sublyrata*	1	2.18 ± 0.12	5.5	2.10	2.27	1068	1.092	5.81	2n=2*x*	*Scrophularia, Scrophularia*
JC06	*Scrophularia sublyrata*	2	2.18 ± 0.11	5.1	2.10	2.26	1067	1.091	4.52	2n=2*x*	*Scrophularia, Scrophularia*

The values are given as mean and standard deviation of the mean of the holoploid genome size (2C, pg) of individuals of each species. For each collection, the ID code and population number (see Appendix 1 for details), the DNA range (Min., minimum genome size; Max., maximum genome size), the monoploid genome size (1C*x*) in Mbp and in mass values (pg), the mean coefficient of variation (CV, %) of G_0/G_1 peaks and the supposed ploidy level are also given. The section and subsection of each *taxon* is also provided. [1] 1 pg = 978 Mbp (Doležel et al. 2003).

Table 7. Nuclear DNA content estimations in the studied *taxa* of *Verbascum*.

ID code	*Taxon*	Pop.	Genome size (2C, pg)				Genome size (1Cx, Mbp)[1]	Genome size (1Cx, pg)	FL CV (%)	Ploidy level	Section, Subsection
			Mean ± SD	CV (%)	Min.	Max.					
MC07	*Verbascum levanticum*	1	0.75 ± 0.02	2.9%	0.74	0.78	368	0.38	5.57	2n=2*x*	*Verbascum, Singuliflora*
MC85	*Verbascum litigiosum*	1	0.76 ± 0.03	4.2%	0.72	0.78	370	0.38	3.48	2n=2*x*	*Verbascum, Verbascum*
MC59	*Verbascum pulverulentum*	1	0.78 ± 0.02	2.2%	0.76	0.80	383	0.39	4.15	2n=2*x*	*Verbascum, Verbascum*
MC36	*Verbascum simplex*	1	0.76 ± 0.02	2.0%	0.75	0.78	373	0.38	3.38	2n=2*x*	*Verbascum, Verbascum*
MC37	*Verbascum simplex*	2	0.76 ± 0.02	2.5%	0.74	0.78	370	0.38	3.03	2n=2*x*	*Verbascum, Verbascum*
MC68	*Verbascum simplex*	3	0.74 ± 0.03	3.8%	0.72	0.77	364	0.37	2.65	2n=2*x*	*Verbascum, Verbascum*
MC100	*Verbascum simplex*	4	0.72 ± 0.03	4.0%	0.70	0.76	354	0.36	4.54	2n=2*x*	*Verbascum, Verbascum*
MC31	*Verbascum sinuatum*	1	0.82 ± 0.01	1.0%	0.82	0.83	403	0.41	4.57	2n=2*x*	*Verbascum, Verbascum*
MC57	*Verbascum sinuatum*	2	0.80 ± 0.01	1.3%	0.79	0.81	391	0.40	4.04	2n=2*x*	*Verbascum, Verbascum*
MC60	*Verbascum sinuatum*	3	0.78 ± 0.05	6.8%	0.74	0.81	380	0.39	3.77	2n=2*x*	*Verbascum, Verbascum*
MC73	*Verbascum sinuatum*	4	0.76 ± 0.03	4.3%	0.72	0.78	370	0.38	3.88	2n=2*x*	*Verbascum, Verbascum*
MNHN35	*Verbascum sinuatum*	5	0.74 ± 0.00	0.2%	0.74	0.75	364	0.37	4.77	2n=2*x*	*Verbascum, Verbascum*
ISAUTL9	*Verbascum sinuatum*	6	0.75 ± 0.02	3.1%	0.73	0.77	366	0.37	9.00	2n=2*x*	*Verbascum, Verbascum*
MC60A	*Verbascum virgatum*	1	1.44 ± 0.02	1.6%	1.41	1.46	703	0.36	2.43	2n=4*x*	*Verbascum, Verbascum*
MNHN36	*Verbascum virgatum*	2	1.41				692	0.35	3.12	2n=4*x*	*Verbascum, Verbascum*

The values are given as mean and standard deviation of the mean of the holoploid genome size (2C, pg) of individuals of each species. For each collection, the ID code and population number (see Appendix 1 for details), the DNA range (Min., minimum genome size; Max., maximum genome size), the monoploid genome size (1C*x*) in Mbp and in mass values (pg), the mean coefficient of variation (CV, %) of G_0/G_1 peaks and the supposed ploidy level are also given. The section and subsection of each *taxon* is also provided. [1] 1 pg = 978 Mbp (Doležel et al., 2003).

Table 8. Nuclear DNA content estimations in the studied *taxa* of *Veronica*

ID code	*Taxon*	Pop.	Genome size (2C, pg)				Genome size (1Cx, Mbp)[1]	Genome size (1Cx, pg)	FL CV (%)	Ploidy level	Section, Subsection
			Mean ± SD	CV (%)	Min.	Max.					
SC21	*Veronica acinifolia*	1	1.24 ± 0.01	0.7	1.23	1.25	608	0.622	3.73	2n=2*x*	*Pocilla, Acinifoliae*
MC05	*Veronica arvensis*	1	0.92 ± 0.00	0.4	0.92	0.93	451	0.461	3.59	2n=2*x*	*Pocilla, Allsinebe*
MC09	*Veronica arvensis*	2	0.91 ± 0.01	0.9	0.90	0.92	447	0.457	4.20	2n=2*x*	*Pocilla, Allsinebe*
MC25	*Veronica arvensis*	3	0.90 ± 0.02	2.6	0.88	0.93	442	0.452	3.98	2n=2*x*	*Pocilla, Allsinebe*
MC11	*Veronica persica*	1	1.45 ± 0.01	0.7	1.43	1.46	354	0.361	6.20	2n=4*x*	*Pocilla, Alsinoides*
MC39	*Veronica persica*	2	1.39 ± 0.04	2.8	1.36	1.43	339	0.347	6.95	2n=4*x*	*Pocilla, Alsinoides*
MC40	*Veronica persica*	3	1.39 ± 0.03	2.1	1.37	1.42	340	0.347	3.55	2n=4*x*	*Pocilla, Alsinoides*
MC41	*Veronica persica*	4	1.38 ± 0.00	0.1	1.38	1.38	336	0.344	4.05	2n=4*x*	*Pocilla, Alsinoides*
MC43	*Veronica persica*	5	1.41 ± 0.05	3.2	1.37	1.46	346	0.354	3.56	2n=4*x*	*Pocilla, Alsinoides*
MC55	*Veronica persica*	6	1.40 ± 0.03	1.9	1.37	1.43	341	0.349	3.07	2n=4*x*	*Pocilla, Alsinoides*
MC61	*Veronica persica*	7	1.41 ± 0.03	2.4	1.37	1.43	344	0.352	3.33	2n=4*x*	*Pocilla, Alsinoides*
MC64	*Veronica polita*	1	0.78 ± 0.01	1.4	0.76	0.78	380	0.389	4.28	2n=2*x*	*Pocilla, Alsinoides*
MC04	*Veronica polita*	2	0.77 ± 0.01	1.7	0.76	0.78	377	0.385	5.09	2n=2*x*	*Pocilla, Alsinoides*
MC03	*Veronica hederifolia*	1	4.16 ± 0.08	2.0	4.10	4.25	678	0.693	2.84	2n=6*x*	*Pocilla, Cymbalariae*
JL07	*Veronica peregrina* subsp. *peregrina*	1	1.96 ± 0.06	2.9	1.90	2.01	319	0.327	4.02	2n=6*x*	*Pocilla, Peregrinae*
MC102	*Veronica chamaedrys* subsp. *chamaedrys*	1	3.72 ± 0.02	0.5	3.70	3.75	607	0.620	3.70	2n=6*x*	*Veronica, Multiflorae*
COI31	*Veronica micrantha*	1	2.14 ± 0.04	1.7	2.11	2.19	525	0.537	6.20	2n=4*x*	*Veronica, Multiflorae*
MC23	*Veronica officinalis*	1	2.18 ± 0.01	0.3	2.17	2.18	533	0.545	2.92	2n=4*x*	*Veronica, Veronica*
JC05	*Veronica officinalis*	2	2.14 ± 0.01	0.4	2.13	2.15	522	0.534	3.34	2n=4*x*	*Veronica, Veronica*
MC95	*Veronica officinalis*	3	2.03 ± 0.05	2.3	1.99	2.08	496	0.508	4.86	2n=4*x*	*Veronica, Veronica*
MC107	*Veronica officinalis*	4	2.08 ± 0.03	1.3	2.06	2.11	507	0.520	4.76	2n=4*x*	*Veronica, Veronica*

The values are given as mean and standard deviation of the mean of the holoploid genome size (2C, pg) of individuals of each species. For each collection, the ID code and population number (see Appendix 1 for details), the DNA range (Min., minimum genome size; Max., maximum genome size), the monoploid genome size (1C*x*) in Mbp and in mass values (pg), the mean coefficient of variation (CV, %) of G_0/G_1 peaks and the supposed ploidy level are also given. The section and subsection of each *taxon* is also provided. [1] 1 pg = 978 Mbp (Doležel et al., 2003)

Chapter 4
DISCUSSION

The amount of DNA per chromosome set is known to be a fairly constant characteristic of a species. While we can pinpoint minute changes in DNA sequences, identify chemical modifications of nuclear bases, and link these changes to the phenotype, we continue to be puzzled by the large variation in the size of the genome itself, of which there seems to be no rational explanation (Greilhuber et al. 2010). Still, since the past decade an increasing interest on genome size studies and its significance has been observed, with many studies focused on using genome size as a taxonomic marker and on finding correlations between ecological and environmental variables and this character. However, there are still many families being neglected such as Scrophulariaceae, for which the present study contributed with more data than the available up to date, with 55 of 59 species (86 %) being new estimates of genome size. Furthermore, due to the importance of polyploidy events on the genesis of new entities, it is important to evaluate how common these events are in nature, and, in the particular case of Scrophulariaceae, how often it may be contributing for the origin of new species in Iberian Peninsula. The detailed bibliographic analyses of polyploidy incidence in this family seemed to point out that at least some *taxa* could present different cytotypes (Appendix 2). However, the absence of more than one cytotype in all the analysed species revealed that polyploidy apparently is not among the main mechanisms of speciation in Scrophulariaceae, currently at least in this region. On the other hand, genome size confirmed to be an excellent tool in species delimitation in many Scrophulariaceae genera.

After molecular studies using DNA sequences of plastid genes, the genera belonging to Scrophulariaceae *sensu latum* were reorganized into 6 different families, including Scrophulariaceae *sensu stricto* (Olmstead et al. 2011). A comparison of genome size taking in consideration this new classification did not reveal any pattern. This result was already expected, as genome size estimations obtained in Scrophulariaceae *s.l.* fell almost exclusively in the very small and small genome size categories (Leitch et al. 1998), presenting a relatively low variation – if *Melampyrum pratense* subsp. *latifolium* is excluded from the analysis (the only species with intermediate genome size), only a 8-fold variation was observed.

As already observed in many genera (*e.g.*, *Helleborus* spp., Zonneveld et al. 2003) genome size can be used as an extra taxonomic character for discriminating between closely related *taxa*. Species belonging to *Bartsia*, *Nothobartsia* and *Parentucellia* share a close evolutionary history and some morphological similarities that sometimes may lead to misidentifications. For example, *Nothobarsia asperrima* was formerly included in the *Bartsia* genus as *Bartsia asperrima* (Benedi et al. 2009). The same

situation repeats with species belonging to *Odontites* and *Odontitella*, with *Odontitella virgatum* being previously included in the former genus (Benedi et al. 2009). As expected by the dissimilar chromosome numbers present in each species of these genera, all the analysed species had different non-overlapping genome sizes, and thus in case any doubt would arise in species identification, using genome estimates the assignment to a taxonomomic category would be straightforward. In a similar study, Loureiro and co-authors, were able to distinghish two genera of Ulmaceae, *Ulmus* and *Celtis* (Loureiro et al. 2007c).

A survey of the Plant DNA C-values database (Bennett and Leitch 2010) revealed a high incidence of intra-generic variation in genome size in species with the same number of chromosomes. At least two-fold variation in monoploid genome size is recorded for more than one third of the genera for which there is sufficient coverage of homoploid species (Suda et al. 2006). Genera *Bulnesia*, *Crepis*, *Cypripedium*, *Dendrobium*, *Lonicera*, *Oxalis*, *Phalaenopsis*, *Scilla*, *Senecio*, *Sisyrinchium*, *Tradescantia*, *Vaccinium* and *Vicia* are some of the best examples of genera with a large between-species divergence in genome size that is not accompanied by changes in the number of chromosomes. Those genera that already were a focus of deep study include *Petunia* (Mishiba et al. 2000), *Hydrangea* (Cerbah et al 2001), *Artemisia* (Torrel and Vallès 2001), *Cistus* (Ellul et al. 2002), *Elytrigia* (Mahelka et al. 2005) and *Curcuma* (Leong-Škorničková et al. 2007), among others. In the case of Scrophulariaceae, contrasting results were obtained among the studied genera: while in a few (*Anarrhinum*, *Antirrhinum* and *Misopates*), genome size was an unsuitable character for taxonomic purposes, as all the estimates were very homogeneous among species; in the other analyzed genera, genome size could be used for *taxa* delimitation and for analyses of interspecific variation. This information was particularly important in the genera with homoploid *taxa*: *Digitalis*, *Pedicularis* and *Linaria*.

In *Digitalis*, all the analysed species had different genome sizes, and indeed, this data supports recent taxonomic changes in this genus: traditionally, *Digitalis mariana* was considered one sub-species of *Digitalis purpurea* and has been recently elevated to the species level (Benedi et al. 2009). Indeed, this new species presents a genome size significantly lower than that of *Digitalis purpurea* subsp. *purpurea*. It will be very interesting to apply FCM to all the species in the genus and evaluate if it continues to be possible to discriminate these homoploid *taxa* using genome size.

In the case of the analysis of the two subspecies of *Pedicularis sylvatica*, as observed in *Crepis foetida* (Dimitrová et al. 1999) and in *Festuca ampla* (Loureiro et al. 2007a), it was possible to separate both subspecies using the information of genome

size. Still, the close proximity of the estimates correlates well with taxonomic promixity of both subspecies. So far, *P. sylvatica* subsp. *sylvatica* was not described for Portugal, so our report in the National Park of Peneda Gerês is a new citation for this country.

In *Linaria*, with exception of *L. triornitophora*, who presented a higher genome size value, all the other species presented more similar genome sizes; still, due to the high quality of the obtained estimates, it was possible to use this character to separate some *taxa*. However, two commonly confused *taxa*, *L. polygalifolia* subsp. *polygalifolia* and *L. supina*, presented the same genome size and thus, unfortunately, could not be distinguished using this character. A rough analysis considering the subgeneric level, seems to point out that members of section *Pelisserianae* present the highest values of genome size, while those from section *Versicolores* present the lowest. Still, this can be do to the reduced number of species analysed in those sections, as evident by the larger heterogeneity in genome size observed in section *Supinae*, the section to which most of the analysed species belong. Previous studies in the literature support this type of sectional analysis. For example, genome size also supports the subgeneric division in the genus *Equisetum*: *taxa* from subgenus *Equisetum* have significantly smaller values than their counterparts from subgenus *Hippochaete* (Obermayer et al. 2002). Similarly, a nuclear DNA content analysis provided information on the sectional classification of the genus *Taraxacum* (Záveský et al. 2005).

In *Veronica*, *Scrophularia* and *Verbascum*, most of the observed differences in genome size were related with different numbers of chromosomes. Still, considering that obtaining good microscopic plates for counting the number of chromosomes in all the analysed species would take a long time, the value of genome size estimates is undeniable also in these cases.

Using genome size, it was possible to distinguish all the analysed *taxa* of *Veronica*, with exception of *V. micrantha* and *V. officinalis*. In a comparison with the only genome size study focused in this genus, some of our estimates are very similar to those of Albach and Greilhuber (2004) (e.g., *V. peregrina* subsp. *peregrina* with 1.93 pg/2C in this thesis vs. 1.90 pg/2C in the literature), while others are clearly different (e.g., *V. chamaedrys* subsp. *chamaedrys* with 3.72 pg/2C in this thesis vs. 2.98 pg/2C in the literature, and *V. arvensis* with 0.91 pg/2C in this thesis vs. 0.66 pg/2C in the literature). Some of these differences could be easily justified by different ploidy levels, as is possibly the case of *V. chamaedrys* and *V. hederifolia* where hexaploidy was assumed in our case instead of tetraploidy (Albach and Greilhuber 2004). Still, in the case of *V. arvensis* the large difference that we observed may be related to the use of different techniques and methodologies. Indeed most of the estimates reported by Albach and

Greilhuber were obtained using Feulgen densitometry, including that of *V. arvensis*. Despite that Doležel et al. (1998) showed a close agreement between both methods, there are numerous cases in the literature where estimates obtained using both techniques do not correspond. For example, Loureiro et al. (2007) using FCM obtained a 2C value of 5.08 pg DNA for *Coriandrum sativum*, while Das and Mallick (1989) using Feulgen microdensitometry obtained 2C values ranging between 7.65 and 9.55 pg/2C. These differences are mostly related with the many critical points of the Feulgen technique (e.g., fixation, slide preparation and storage, acid hydrolysis), which are not always followed and that may influence the obtained estimations (Greilhuber, 1988). Still, particularities of the FCM methodology, as the use of different reference standards, sample preparation and staining protocols (Doležel et al 1998), may also contribute for these differences.

Following the analysis of monoploid genome size variation and the linear regression between chromosome numbers and genome size, besides the different DNA ploidy level assumed for *V. chamaedrys* and *V. hederifolia*, it seems that the the analysed individuals of *V. micrantha* and *V. officinalis* are tetraploid and not diploid, as assumed in the literature. Thus, the sampled populations present 32 and 36 chromosomes, respectively. As these are first estimates of genome size, no information on this parameter is available in the literature that could be used to certify this assumption. Still, in the case of *V. officinalis*, there are some previous reports of 36 chromosomes despite two base chromosome numbers, 9 and 18, are reported (Benedi et al. 2009), indicating some confusion to what ploidy level the set of 36 chromosomes corresponds. Therefore this hypothesis needs to be confirmed in the future using chromosome counts and fluorescence in situ hybridization (FISH). Similar results are documented in other families. For example, in *Festuca*, Al-Bermani et al. (1992) attributed the octaploid level to *F. rothmaleri*, while, more recently, Loureiro et al. (2007) detected hexaploids in this species.

In *Scrophularia*, several species had apparently different genome sizes, but those differences revealed to be not statistically significant. The use of a non-parametric statistical test due to problems in achieving homocedasticity (even after data transformation) can explain the lack of statistical differences among *taxa*. In this genus, the species with the highest number of chromosomes, *S. auriculata* subsp. *auriculata*, is not the one with the higher value of genome size. Considering the number of chromosomes that this species presents (78-88 chromosomes) it is certain that it suffered from several polyploidy events in the past and, as happened in other species (e.g., *Nicotiana* sp., Leitch et al. 2008), these phenomena may have been

accompanied by genome downsizing. It is assumed that DNA loss during polyploidization may be a selection mechanism to lessen genetic instability or the phenotypic effects of having a larger nucleus and cell size (Leitch et al. 2008). A rough analysis of genome size variation among sections revealed that species from Section *Caninae* presented, in general, smaller genome sizes than those of Section *Scrophularia*. Still, *S. nodosa* from section Scrophularia presented the smallest genome size among the analysed *taxa* of *Scrophularia*, being in complete disagreement with this postulate.

Finally, in *Verbascum*, if we exclude *V. virgatum* who is tetraploid and consequently presents approximately the double value of genome size of the remaining species, the other *taxa* presented very similar genome sizes. Still, all these species present different chromosome numbers (*i.e.*, 30, 32, 36 chromosomes; Benedi et al. 2009). This may be due to a phenomenon called disploidy, i.e., the increase or decrease of one, or a few chromosomes. The decrease in chromosome numbers appears not to be unusual (Martel et al. 2004, Hidalgo et al. 2007) and may be due to the fusion of two or more chromosomes. In principle, this would not affect the genome size in any way. Based on chromosome number variation, descendant disploidy has been suggested for several genera of Iridaceae (Goldblatt and Takei 1997). For example, in *Iris* subgenus *Xiphium*, it was proposed that if the ancestral base number was x = 9, and *I. boissieri* (n = 18) represented a polyploidy event, descending disploidy may explain the remaining chromosome numbers (n = 17, 16, 15, 14). Similarly, in *Verbascum*, the same phenomenon may explain a decrease in the number of chromosomes from 36, to the remaining chromosomes numbers that are reported in the literature (2n = 30, 32 and 34 chromosomes), without variation in genome size. Molecular cytogenetic techniques as FISH could help to solve this question and should be used in the future.

The analysis of intraspecific variation revealed some variation in genome size among individuals of the same species, both between and within populations. While some argue for a large plasticity of nuclear genome, others claim for a more stable genome size within species. In reality, the growing number of reports that did not confirm the intraspecific variation reported in original publications [see list of “blunder-killing” papers of Suda (2004) and the review by Greilhuber (2005)] has shifted the pendulum towards the stability side, without eliminating the possibility of its occurrence. Actually, in recent years several reports that followed best practices confirmed the existence of this phenomenon (see Šmarda and Bureš 2010 for a review). In the case there is a true intraspecific variation, chromosomal differences (aneuploidy and supernumerary B-chromosomes,) and polymorphisms in A chromosomes

(heterochromatic knobs and differential deletion of transposable element remnants) (Gregory 2005a), may explain the differences that were reported. In particular, it is worth highlighting the differences observed in the genome size estimates among individuals of *Digitalis purpurea* subsp. *purpurea*. In here, three groups differing of about 0.11 pg were observed. In the literature, two chromosome numbers, 48 and 56, are known (Benedi et al. 2009). Furthermore, the possibility of presenting B chromosomes is documented for this species (Regnart 1934). Alltogether a combination of both these events may support the variation observed in this subspecies, similarly to what was reported by Sharbel et al. (2004) in *Boechera holboellii*. To fully confirm the occurrence of this phenomenon, chromosome counts of this subspecies should be prepared in the future.

In conclusion, this work contributed with important background knowledge on genome size variation and polyploidy incidence in the Iberian Peninsula Scrophulariaceae. Despite the results on polyploidy incidence were discouraging, with no multiple cytotypes being detected for any *taxa*, genome size results confirmed the high importance this character may have in species delimitation. Indeed, regarding genome size evolution, many doors were open, with interesting phenomena, as intraspecific variation of genome size and disploidy being detected. Future studies should focus in studying the genome size of the remaining *taxa* of Scrophulariaceae, while broadening sampling to central and eastern Iberian Peninsula.

Chapter 5
REFERENCES

Adams KL, Wendel J F (2005) Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology* **8**:135-141.

Albach DC, Greilhuber J (2004) Genome size variation and evolution in *Veronica. Annals of Botany* **94**:897-911.

Al-Bermani AK, Catalan P, Stace CA (1992) A new circumscription of *Festuca trichophylla* (Gaudin) K. Richter (Gramineae). *Anales del Jardín Botánico de Madrid* **50**:209-220.

Angiosperm Phylogeny Group (1998) An ordinal classification of the families of flowering plants. *Annals of the Missouri Botanical Garden* **85**:531–553.

Angiosperm Phylogeny Group II (2003) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* **141:**399–436.

Baack EJ (2004) Cytotype segregation on regional and microgeographic scales in snow buttercups (*Ranunculus adoneus*: Ranunculaceae). *American Journal of Botany* **91**:1783-1788.

Beaulieu JM, Moles AT, Leitch IJ, Bennett MD, Dickie JB, Knight CA (2007) Correlated evolution of genome size and seed mass. *New Phytologist* **173**:422-437.

Beaulieu JM, Leitch IJ, Patel S, Pendharkar A, Knight CA (2008) Genome size is a strong predictor of cell size and stomatal density in angiosperms. *New Phytologist* **179**:975-986.

Benedí C, Rico E, Güemes J, Herrero A (200X) Plantaginaceae-Scrophulariaceae. In: Castroviejo et al. (eds.). Flora Iberica Vol. 13, Real Jardín Botánico, Madrid, pp. 44-434

Bennert W, Lubienski M, Körner S, Steinberg M (2005) Triplody in *Equisetum* subgenus *Hippochaete* (Equisetaceae, Pteridophyta). *Annals of Botany* **95**:807–15.

Bennett MD (1971) Duration of Meiosis*. Proceedings of the Royal Society of London Series B-Biological Sciences* **178**:277-&.

Bennett MD (1972) Nuclear DNA content and minimum generation time in herbaceous plants. *Proceedings of the Royal Society B* **181**:109–135.

Bennett MD (1977) Time and duration of meiosis. *Philosophical Transactions of the Royal Society* B **277**:201–226.

Bennett MD, Smith JB (1976) Nuclear DNA amounts in angiosperms. *Philosophical of the Royal Society B* **274**:228-274.

Bennett MD, Leitch IJ (2005) Nuclear DNA amounts in angiosperms: progress, problems and prospects. *Annals of Botany* **95**:45–90.

Bennett MD, Leitch IJ (2010) Angiosperm DNA *C*-values database (release 7.0, Dec.2010)

Bennett MD, Leitch IJ, Hanson L (1998) DNA amounts in two samples of angiosperm weeds. *Annals of Botany* **82**:121–134.

Bennetzen JL, Ma JX, Devos K (2005) Mechanisms of recent genome size variation in flowering plants. *Annals of Botany* **95**:127–132.

Bremer B, Olmstead RG, Struwe L, Sweere JA (1994) *RbcL* sequences support exclusion of *Retzia, Desfontainia,* and *Nicodemia* (Buddlejaceae) from the Gentianales. *Plant Systematics and Evolution* **190**:213–230.

Brook BW, Sodhi NS, Ng PKL (2003) Catastrophic extinctions follow deforestation in Singapore. *Nature* **424**:420-423.

Bureš P, Tichý L, Wang YF, Bartoš J (2003) Occurrence of *Polypodium* x *mantoniae* and new localities for *P. interjectum* in the Czech Republic confirmed using flow cytometry. *Preslia* **75**:293–310.

Bureš P, Wang YF, Horová L, Suda J (2004) Genome size variation in Central European species of *Cirsium* (Compositae) and their natural hybrids. *Annals of Botany* **94**:353–363.

Burton TL, Husband BC (1999) Population cytotype structure in the polyploid *Galax urceolata* (Diapensiaceae). *Heredity* **82**:381-390.

Burton TL, Husband BC (2001) Fecundity and offspring ploidy in matings among diploid, triploid and tetraploid *Chamerion angustifolium* (Onagraceae): consequences for tetraploid establishment. *Heredity* **87**:573–82.

Castro S, Münzbergová Z, Raabová J, Loureiro J (2011) Breeding barriers at a diploid–hexaploid contact zone in *Aster amellus*. *Evolutionary Ecology* 25:795-814.

Cerbah M, Mortreau E, Brown S, Siljak-Yakovlev S, Bertrand H, Lambert C (2001) Genome size variation and species relationships in the genus *Hydrangea*. *Theoretical and Applied Genetics* **103**:45-51.

Chapin FS (2003) Effects of plant traits on ecosystem and regional processes: a conceptual framework for predicting the consequences of global change. *Annals of Botany* **91**:455-463.

Côrte-Real M, Sansonetty F, Ludovico P, Prudêncio C, Rodrigues F, Fortuna M, Sousa M, Silva M, Leão C (2002) Contributos da citologia analítica para estudos de biologia de leveduras. *Boletim de Biotecnologia* **71**:19-33.

Dimitrová D, Ebert I, Greilhuber J, Kozhuharov S (1999) Karyotype constancy and genome size variation in Bulgarian *Crepis foetida s.l.* (Asteraceae). *Plant Systematics and Evolution* **217**:245-257.

Doležel J (1997) Applications of Flow Cytometry for the study of plant genomes. *Journal of Applied Genetics* **38**:285-302.

Doležel J, Bartos J (2005) Plant DNA flow cytometry and estimation of nuclear genome size. *Annals of Botany* **95:**99-110.

Dolezel J, Sgorbati S, Lucretti S (1992) Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiologia Plantarum* **85**:625-631.

Dolezel J, Dolezelová M, Novák F (1994) Flow cytometric estimation of nuclear DNA amount in diploid bananas (*Musa acuminata* and *M. balbisiana*). *Biologia Plantarum* **36:**351-357.

Doležel J, Greilhuber J, Suda J (2007) Flow cytometry with plants: an overview. In: Doležel J, Greilhuber J, Suda J (eds), Flow cytometry with plant cells, Wiley VCH, Weinheim pp. 41–65.

Doležel J, Bartoš J, Voglmayr H, Greilhuber J (2003) Nuclear DNA content and genome size of trout and human. *Cytometry Part A* **51A**:127–128.

Doležel J, Geilhuber J, Lucretti S, Meister A, Lysák M, Nardi L, Obermayer R (1998) Plant genome size estimation by flow cytometry: inter-laboratory comparison. *Annals of Botany* **82**:17-26.

Ekrt L, Trávníček P, Jarolímová V, Vít P, Urfus T (2009) Genome size and morphology of the *Dryopteris affinis* group in Central Europe. *Preslia* **81**:261–280.

Ellul P, Boscaiu M, Vicente O, Moreno V, Rossello JA (2002) Intra- and interspecific variation in DNA content in *Cistus* (Cistaceae). *Annals of Botany* **90**:345-351.

Faegri K, van der Pijl L (1979) The Principles of Pollination Ecology. Pergamon Press, Oxford.

Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E (1983) Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science* **220**:1049-1051.

Goldblatt P, Takei M (1997) Chromosome cytology of Iridaceae – patterns of variation, determination of ancestral base numbers, and modes of karyotype change. *Annals of the Missouri Botanical Garden* **4**:285–304.

Greenlee JK, Rai KS, Floyd AD (1984) Intraspecific variation in nuclear DNA content in *Collinsia verna* Nutt (Scrophulariaceae). *Heredity* **52**:235-242.

Gregory TR (2001) Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. *Biological Reviews of the Cambridge Philosophical Society* **76**:65–101.

Gregory TR (2005a) The C-value enigma in plants and animals: a review of parallels and an appeal for partnership. *Annals.of Botany* **95**:133–46.

Gregory TR (2005b) Synergy between sequence and size in large-scale genomics. *Nature Reviews Genetics* **6**:699–708

Greilhuber J (1988) Self tanning - a new and important source of stoichiometric error in cytophotometric determination of nuclear DNA content in plants. *Plant Systematics and Evolution* **158**:87-96.

Greilhuber J (1998) Intraspecific variation in genome size: a critical reassessment. *Annals of Botany* **82**:27–35.

Greilhuber J (2005) Intraspecific variation in genome size in angiosperms: identifying its existence. *Annals of Botany* **95**:91–98.

Greilhuber J, Speta F (1985) Geographical variation of genome size at low taxonomic levels in the *Scilla bifolia*, alliance (Hyacinthaceae). *Flora (Jena)* **176**:431–438.

Greilhuber J, Temsch E, Loureiro J (2007) Nuclear DNA content measurement. In: Dolezel J, Greilhuber J, Suda J (eds). Flow cytometry with plant cells. Weinheim: Wiley-VCH Verlag, pp 67-101.

Greilhuber J, Doležel J, Lysák MA, Bennett MD (2005) The origin, evolution and proposed stabilization of the terms 'genome size' and 'C-value' to describe nuclear DNA contents. *Annals of Botany* **95**:255–60

Greilhuber J, Borsch T, Müller K, Worberg A, Porembski S, Barthlott W (2006) Smallest angiosperm genomes found in Lentibulariaceae with chromosomes of bacterial size. *Plant Biology* **8**: 770-777

Greilhuber J, Dolezel J, Leitch I, Loureiro J, Suda J (2010) Genome size. *Journal of Botany* **Volume 2010**: 4 pages

Habitats Directive ofNATURA2000—Web 1: http://www.nature.cz/publik_syst2/files08/habitats%20directive_official%20text.pdf

Hidalgo O, Garcia-Jacas N, Garnatje T, Susanna A, Siljak-Yakovlev S (2007) Karyological evolution in *Rhaponticum* Vaill. (Asteraceae, Cardueae) and related genera. *Botanical Journal of the Linnean Society* **53**:193–201.

Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman AH (2000) Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. *Proceedings of the National Academy of Science USA* **97**:6603–6607.

Kikuchi S, Tanaka H, Shiba T, Mii M, Tsujimoto H (2006) Genome size, karyotype, meiosis and a novel extra chromosome in *Torenia fournieri*, *T. baillonii* and their hybrid. *Chromosome Research* **14**:665-672.

Knight CA, Molinari NA, Petrov DA (2005) The large genome constraint hypothesis: evolution, ecology and phenotype. *Annals of Botany* **95**:177–190.

Kron P, Suda J, Husband BC (2007) Applications of flow cytometry to evolutionary and population biology. *Annual Reviews of Ecology, Evolution, and Systematics* **38**:847-76.

Lampert KP, Lamatsch DK, Epplen JT, Schartl M (2005) Evidence for a monophyletic origin of triploid clones of the Amazon molly, *Poecilia formosa*. *Evolution* **59**:881–89.

Leitch IJ, Chase MW, Bennett MD (1998) Phylogenetic analysis of DNA C-values provides evidence for a small ancestral genome size in flowering plants. *Annals of Botany* **82**:85-94.

Leitch IJ, Hanson L, Lim KY, Kovarik A, Chase MW, Clarkson JJ, Leitch AR (2008) The ups and downs of genome size evolution in polyploid species of *Nicotiana* (Solanaceae). *Annals of Botany* **101**:805–814.

Leong-Škorničková J, Šída O, Jarolímová V, Sabu M, Fér T, Trávníček P, Suda J (2007) Chromosome numbers and genome size variation in Indian species of *Curcuma* (Zingiberaceae). *Annals of Botany* **100**:505–526.

Levin DA (2002) The role of chromosomal change in plant evolution. Oxford University Press, New York.

Lim KY, Kovarik A, Matyasek R, Chase MW, Knapp S, McCarthy E, Clarkson JJ, Leitch AR (2006) Comparative genomics and repetitive sequence divergence in the species of diploid *Nicotiana* section *Alatae*. *Plant Journal* **48**:907–919.

Lista Roja de Flora Vascular Española (valoración según categorías UICN). *CONSERVACIÓN VEGETAL* **6** (extra):11-38.

Lobo JR, Lumaret JP, Jay-Robert P (2001) Diversity, distinctiveness and conservation status of the Mediterranean coastal dung beetle assemblage in the Regional Natural Park of the Camargue (France). *Diversity and Distributions* **7**:257-270.

Loureiro J (2007) Flow cytometric approaches to study genomes. PhD Thesis, University of Aveiro.

Loureiro J, Rodriguez E, Dolezel J, Santos C (2006) Flow cytometric and microscopic analysis of the effect of tannic acid on plant nuclei and estimation of DNA content. *Annals of Botany* **98**:515-527.

Loureiro J, Kopecky D, Castro S, Santos C, Silveira P (2007a) Flow cytometric and cytogenetic of Iberian Peninsula *Festuca* spp. *Plant Systematics and Evolution* **269**:89-105.

Loureiro J, Rodriguez E, Dolezel J, Santos C (2007b) Two new nuclear isolation buffers for plant DNA flow cytometry: A test with 37 species. *Annals of Botany* **100**:875-888.

Loureiro J, Rodriguez E, Gomes A, Santos C (2007c) Genome size estimations on *Ulmus minor* Mill., *Ulmus glabra* Huds., and *Celtis australis* L. using flow cytometry. *Plant Biology* **9**:541-544.

Loureiro J, Trávnícek P, Rauchová J, Urfus T, Vit P, Štech M, Castro S, Suda J (2010) The use of flow cytometry in the biosystematics, ecology and population biology of homoploid plants. *Preslia* **82**:3-21.

Lysák M, Dolezel J (1998) Estimation of nuclear DNA content in Sesleria (Poaceae). *Caryologia* **52**:123-132.

Mable BK (2004) 'Why polyploidy is rarer in animals than in plants': myths and mechanisms. *Biological Journal of the Linnean Society* **82**:453–66.

Mahelka V, Suda J, Jarolímová V, Trávníček P, Krahulec F (2005) Genome size discriminates between closely related *taxa Elytrigia repens* and *E. intermedia* (Poaceae: Triticeae) and their hybrid. *Folia Geobotanica* **40**:367–384.

Martel E, Poncet V, Lamy F, Siljak-Yakovlev S, Lejeune B, Sarr A (2004) Chromosome evolution of *Pennisetum* species (Poaceae): implications of ITS phylogeny. *Plant Systematics and Evolution* **249**:139–149.

Martínez J, Vargas P, Luceño M (2010) Evolution of *Iris* subgenus *Xiphium* based on chromosome numbers, FISH of nrDNA (5S, 45S) and *trnL- trnF sequence analysis. Plant Systematics and Evolution* **289**:223-235.

Maxted N, Callimassia MA, Bennett MD (1991) Cytotaxonomic studies of eastern mediterranean *Vicia* species (Leguminosae). *Plant Systematics and Evolution*. **177**:221–234.

Meagher TR and Costich DE (1994) Sexual dimorphism in nuclear DNA content and floral morphology in populations of *Silene latifolia* (Caryophyllaceae). *American Journal of Botany* **81**:1198–1204.

Médail F, Quézel P (1997) Hot-spots analysis for conservation of plant biodiversity in the Mediterranean Basin. *Annals of the Missouri Botanical Garden* **84**:112–127.

Mishiba K, Ando T, Mii M, Watanabe H, Kokubun H, Hashimoto G, Marchesi E (2000) Nuclear DNA content as an index character discriminating *taxa* in the genus *Petunia* sensu ussieu (*Solanaceae*). *Annals of Botany* **85**:665–673.

Morton JK (1993) Chromosome numbers and polyploidy in the flora of Cameroons Mountain. *Opera Botanica* **121:**159– 172.

Mowforth MAG (1986) Variation in nuclear DNA amounts in flowering plants: an ecological analysis. PhD Thesis, University of Sheffield.

Myers N (1989) Extinction rates past and present. *Bioscience* **39**:39-41.

Myers N, Mittermeier RA, Mittermeier CG, Fonsaeca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* **403**:853-858.

Nagl W, Fusenig HP (1979) Types of chromatin organization in plant nuclei. *Plant Systematics and Evolution* **Suppl. 2**:221-233.

Noirot M, Barre P, Louarn J, Duperray C, Hamon S (2002) Consequences of stoichiometric error on nuclear DNA content evaluation in *Coffea liberica* var. *dewevrei* using DAPI and propidium iodide. *Annals of Botany* **89**:385-389.

Novacek MJ, Cleland EE (2001) The current biodiversity extinction event: scenarios for mitigation and recovery. *Proceedings of the National Academy of Sciences of the United States of America* **98**:5466-5470.

Obermayer R, Leitch IJ, Hanson L, Bennett MD (2002) Nuclear DNA C-values in 30 species double the familial representation in pteridophytes. *Annals. of Botany* **90**:209–217.

Olmstead RG (2002) Whatever happened to the Scrophulariaceae? *Fremontia* **3**:13-22.

Olmstead RG, Bremer B, Scott K, Pallmer JD (1993) A parsimony analysis of the Asteridae sensu lato based on *rbcL* sequences. *Annals of the Missouri Botanical Garden* **80**:700–722.

Olmstead RG, de Pamphilis CW, Wolfe AF, Young ND, Elisons WJ, Reeves PA (2001) Disintegration of the Scrophulariaceae. *American Journal of Botany* **88**:348-361.

Otto SP, Whitton J (2000) Polyploidy: Incidence and evolution. *Annual Reviews of Genetics* **34**:401-437.

Pecinka A, Suchánková P, Lysák MA, Trávníček B, Doležel J (2006) Nuclear DNA content variation among central European *Koeleria taxa*. *Annals of Botany* **98**:117-22.

Pellicer J, Fay MF, Leitch IJ (2010) The largest eukaryotic genome of them all? *Botanical Journal of the Linnean Society* **164**:10-15.

Pimm SL, Russell GJ, Gittleman JL, Brooks TM (1995) The future of biodiversity. *Science* **269**:347-350.

Price HJ, Johnston JS (1996) Influence of light on DNA content of *Helianthus annuus* Linnaeus. *Proceedings of the National Academy of Sciences of the United States of America* **93**:11264-11267.

Ramsey JR, Schemske DW (1998) Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Reviews of Ecology, Evolution, and Systematics* **29**:467-501.

Ramsey J, Schemske DW (2002) Neopolyploidy in flowering plants. *Annual Review of Ecology, Evolution and Systematics* **33**:589–639.

Raven P (1987) The scope of the plant conservation problem worldwide. In: Bramwell D, Hammaa O, Heywood V, Synge H (eds.) *Botanic gardens of the world conservation strategy*. *London: Academic Press* pp. 19-29.

Regnart HC (1934) Studies of hybrids in the genus *Digitalis*. *Department of Botany and Genetics, Armstrong College, University of Durham, Newcastle-upon-Tyne* pp 145-153

Sala OE, Chapin FS, Armesto JJ, Berlow E, Bloomfield J, Dirzo R, Huber-Sanwald E, Huenneke LF, Jackson RB, Kinzig A, Leemans R, Lodge DM, Mooney HA, Oesterheld M, Poff NL, Sykes MT, Walker BH, Walker M, Wall DH (2000) Global biodiversity scenarios for the year 2100. *Science* **287**:1770-1774.

Sharbel TF, Voigt ML, Mitchell-Olds T, Kantama L, de Jong H (2004) Is the aneuploid chromosome in an apomictic *Boechera holboellii* a genuine B chromosome? *Cytogenetics and Genome Research* **106:**173–183.

Soltis DE, Soltis PS, Tate JA (2003) Advances in the study of polyploidy since plant speciation. *New Phytologist* **161**:173–91.

Stace CA (2000) Cytology and cytogenetics as a fundamental taxonomic resource for the 20th and 21st centuries. *Taxon* **49**:451–77.

Stevens PF (2001 onwards) Angiosperm Phylogeny Website. Version 9, June 2008 [and more or less continuously updated since].

Suda J (2004) An employment of flow cytometry into plant biosystematics. PhD Thesis, Charles University in Prague.

Suda J, Travnicek P (2006) Reliable DNA ploidy determination in dehydrated tissues of vascular plants by DAPI flow cytometry - New prospects for plant research. *Cytometry Part A* **69A**:273-280.

Suda J, Krahulcová A, Travnicek P, Krahulec F (2006) Ploidy level versus DNA ploidy level: an appeal for consistent terminology. *Taxon* **55**:447-450.

Suda J, Krahulcová A Trávníček P, Rosenbaumová R, Peckert T, Krahulec F (2007a) Genome size variation and species relationships in *Hieracium* subgenus *Pilosella* (Asteraceae) as inferred by flow cytometry. *Annals of Botany* **100**:1323–1335.

Suda J, Kron P, Husband BC, Trávníček P (2007b) Flow cytometry and ploidy: applications in plant systematics, ecology and evolutionary biology. In: Doležel J, Greilhuber J, Suda J (eds), Flow cytometry with plant cells, Wiley VCH, Weinheim pp. 103–30

Swift H (1950) The constancy of desoxyribose nucleic acid in plant nuclei. Proceedings *of the National Academy of Sciences* **36**:643–654.

Thieret JW (1967) Supraspecific classification in the Scrophulariaceae: a review. *Sida* **3**:87–106.

Thompson J (1999) Population differentiation in Mediterranean plants: insights into colonization history and the evolution and conservation of the endemic species. *Heredity* **82**:229-236.

Thompson JD, Lumaret R (1992) The evolutionary dynamics of polyploid plants: origins, establishment and persistence. *Trends in Ecology and Evolution* **7**:302–307.

Tiersch TR, Chandler RW, Wachtel SS, Elias S (1989) Reference standards for flow cytometry and application in comparative studies of nuclear DNA content. *Cytometry* **10:**706-710.

Torrell M, Vallès J (2001) Genome size in 21 *Artemisia* L. species (*Asteraceae*, *Anthemideae*): systematic, evolutionary, and ecological implications. *Genome* **44**:231–238.

Trucco F, Tatum T, Robertson KR, Rayburn AL, Tranel PJ (2006) Characterization of waterhemp (*Amaranthus tuberculatus*) × smooth pigweed (*A. hybridus*) F1 hybrids. *Weed Technology* **20**:14–22.

Vargas P, Rosselló JA, Oyama R, Güemes J (2004) Molecular evidence for naturalness of genera in the tribe *Antirrhineae* (Scrophulariaceae) and three independent evolutionary lineages from the New World and the Old. *Plant Systematics and Evolution* **249**:151-172.

Vitte C and Bennetzen JL (2006) Analysis of retrotransposon structural diversity uncovers properties and propensities in angiosperm genome evolution. *Proceedings of the National Academy of Science USA* **103**:17638–17643.

von Wettstein R (1891) Scrophulariaceae. In: Engler A, Prantl K (eds.), Die Natürlichen Pflanzenfamilien, Vol. 4, pp.39–107.

Warren MS, Hill JK, Thomas JA, Asher J, Fox R, Huntley B, Roy DB, Telfer MG, Jeffcoate S, Harding P, Jeffcoate G, Willis SG, Greatorex-Davies JN, Moss D, Thomas CD (2001) Rapid responses of British butterflies to opposing forces of climate and habitat change. *Nature* **414**:65-69.

Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH (2009) The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences* **106**:13875–13879.

Yeater KM, Bollero GA, Bullock DG, Rayburn AL (2004) Flow cytometric analysis for ploidy level differentiation of 45 hairy vetch accessions. *Annals of Applied Biology* **145**:123–127.

Záveský L, Jarolímová V, Štěpánek J (2005) Nuclear DNA content variation within the genus *Taraxacum* (Asteraceae). *Folia Geobotanica* **40:**91–104.

Zonneveld BJM (2001) Nuclear DNA contents of all species of *Helleborus* (Ranunculaceae) discriminate between species and sectional divisions. *Plant Systematics and Evolution* **229**:125-130.

Chapter 6
APPENDICES

Appendix 1. Field collections of Scrophulariaceae *s.l. taxa* performed during the flowering season (March to August) 2010 and 2011 in Portugal and Spain. For each *taxon* an ID code, information on the location and GPS coordinates of the population are given. Voucher specimens are kept in the Herbarium of University of Coimbra (COI).

ID code[1]	*Taxon*	Location	GPS coordinates	
MC67	*Anarrhinum bellidifolium*	PT: Seia, Aldeia da Serra	40°24'51.56"N	7°41'28.30"W
MC65	*Anarrhinum bellidifolium*	PT: Oliveira do Hospital, Caldas da Felgueira	40°29'48.24"N	7°50'14.43"W
MC72	*Anarrhinum bellidifolium*	PT: Arganil, Pardieiros	40°13'30.18"N	7°56'16.30"W
MC15	*Anarrhinum bellidifolium*	PT: Coimbra, Caneiro	40°10'43.57"N	8°19'6.23"W
MC28	*Anarrhinum bellidifolium*	PT: Condeixa-a-Nova, Casmilo	40°03'14.23"N	8°29'56.29"W
SC35	*Anarrhinum bellidifolium*	PT: Bragança, road between Alimonde and Carrazedo	41º47'7.94"N	6º53'20.51"W
MC96	*Anarrhinum bellidifolium*	PT: Serra do Gerês, Borrageiro	41º46'42.27"N	7º 56'17.70"W
SC31	*Anarrhinum bellidifolium*	SP: Galiza, Folgoso de Caurel, Mércurim	42º37'34.74"N	7º10'5.63"W
SC32	*Anarrhinum bellidifolium*	SP: Galiza, Folgoso de Caurel, Seoane de Caurel (Cotelo)	42º38'18.36"N	7º9'3.06"W
JC04	*Anarrhinum bellidifolium*	PT: Serra do Gerês, Borrageiro	41º45'47.05"N	8º7'40.51"W
UPT13	*Anarrhinum bellidifolium*	PT: Tabuaço, Desejosa	-	-
MC33	*Anarrhinum duriminium*	PT: Montalegre, Pitões das Junias, Planalto da Morela	-	-
UPT14	*Anarrhinum duriminium*	PT: Marco de Canaveses, Paços de Gaiolo	-	-
COI1	*Anarrhinum duriminium*	PT: Castro Daire	-	-
MC52	*Anarrhinum longipedicelatum*	PT: Vale de Cambra, Souto Mau	40°46'52.80"N	8°16'50.80"W
MC35	*Anarrhinum longipedicelatum*	PT: Aveiro, Sever do Vouga, Couto de Esteves	40°45'32.63"N	8°18'27.40"W
JP01	*Anarrhinum longipedicelatum*	Unknown locality	-	-
COI4	*Antirrhinum cirrhigerum*	PT: Gala	-	-
COI3	*Antirrhinum graniticum*	PT: Valhelhas, Guarda	-	-
MC24	*Antirrhinum linkianum*	PT: Coimbra, Buçaco	40°22'33.82"N	8°21'50.99"W
MC06	*Antirrhinum linkianum*	PT: Coimbra, Jardim Botânico	40°12'23.06"N	8°25'31.14"W

MC08	*Antirrhinum linkianum*	PT: Coimbra, Santa Clara	40°11'33.83"N	8°25'56.85"W
SC15	*Antirrhinum linkianum*	PT: Setúbal, Comporta	38°22'59.43"N	8°47'51.50"W
JC09	*Antirrhinum linkianum*	PT: Leiria, Nazaré, Sítio	39º36'17.31"N	9º5'3.68"W
JC10	*Antirrhinum linkianum*	PT: Leiria, Nazaré, Praia do Norte	39º36'18.40"N	9º5'4.96"W
ISAUTL3	*Antirrhinum linkianum*	PT: Sintra	-	-
MC70	*Antirrhinum onubense*	PT: Oliveira do Hospital, Avô	40°17'56.95"N	7°54'5.42"W
MC75	*Bartsia trixago*	PT: Serra d'Aires e Candeeiros, Casal das Pias	39°33'32.98"N	8°48'31.46"W
SC42	*Bartsia trixago*	PT: Figueira da Foz	40°12'8.50"N	8°52'50.45"W
MC78	*Bartsia trixago*	PT: São Martinho do Porto	39°30'50.29"N	9° 8'29.47"W
SC19	*Bartsia trixago*	PT: Lisboa, Paredes	38°41'52.80"N	9°22'3.20"W
SC24	*Bartsia trixago*	PT: Cascais, Praia do Guincho	38º43'35.42"N	9º28'27.17"W
UPT17	*Bartsia trixago*	PT: Régua	-	-
MC62	*Chaenorhinum origanifolium*	PT: Serra d'Aires e Candeeiros, Casal das Pias	39°33'35.47"N	8°48'30.76"W
MC38	*Chaenorhinum origanifolium*	PT: Alcobaça, Casal de Vale de Vento	39°28'32.67"N	8°54'30.00"W
MC47	*Chaenorhinum origanifolium*	PT: Serra d'Aires e Candeeiros, Casal de Vale de Ventos	39°27'5.62"N	8°54'39.87"W
MC109	*Cymbalaria muralis* subsp. *muralis*	SP: Galiza, Folgoso de Caurel Samos	42°43'55.65"N	7°19'36.70"W
MC21	*Cymbalaria muralis* subsp. *muralis*	PT: Coimbra, Luso	40°22'42.30"N	8°22'12.10"W
MC02	*Cymbalaria muralis* subsp. *muralis*	PT: Coimbra, Jardim Botânico	40°12'23.61"N	8°25'29.68"W
COI8	*Digitalis mariana* subsp. *mariana*	PT: Numão, castle	-	-
SC34	*Digitalis purpurea* subsp. *purpurea*	PT: Bragança, road between Alimonde and Carrazedo	41º47'7.94"N	6º53'20.51"W
SC39	*Digitalis purpurea* subsp. *purpurea*	PT: Bragança, Carrazeda de Ansiães	41º13'58.64"N	7º19'32.11"W
SC41	*Digitalis purpurea* subsp. *purpurea*	PT: Bragança, Parambos	41º14'16.61"N	7º21'57.69"W
MC69	*Digitalis purpurea* subsp. *purpurea*	PT: Seia, Aldeia da Serra	40°25'10.96"N	7°41'1.10"W
MC20	*Digitalis purpurea* subsp. *purpurea*	PT: Aveiro, Sever do Vouga, Parada	40°46'13.00"N	8°17'35.33"W

MC17	*Digitalis purpurea* subsp. *purpurea*	PT: Coimbra, Vale de Canas	40°12'37.25"N	8°22'33.06"W
MC94	*Digitalis purpurea* subsp. *purpurea*	PT: Serra do Gerês, Borrageiro	41º46'42.27"N	7º 56'17.70"W
SC29	*Digitalis purpurea* subsp. *purpurea*	SP: Galiza, Folgoso de Caurel, Santa Eufemia	42º34'12.67"N	7º11'49.68"W
MC92	*Digitalis purpurea* subsp. *purpurea*	PT: Serra do Gerês, Borrageiro	41º46'04.09"N	8º 06'48.14"W
UPT20	*Digitalis purpurea* subsp. *purpurea*	PT: Cinfães, Ferreiros	-	-
COI9	*Digitalis thapsi*	PT: Seia	-	-
MC105	*Euphrasia minima*	SP: Galiza, Folgoso de Caurel, Visuña	42°35'56.41"N	7° 3'12.60"W
ISAUTL5	*Keckxia spuria* subsp. *integrifolia*	PT: Lisboa, Tapada de Ajuda	-	-
SC28	*Linaria aeruginea* subsp. *aeruginea*	PT: Guarda, Vila Nova de Foz Côa	-	-
MC49	*Linaria amethystea* subsp. *amethystea*	PT: Serra d'Aires e Candeeiro, Casal de Vale de Ventos	39°26'42.07"N	8°54'50.65"W
MC14	*Linaria diffusa*	PT: Coimbra, Caneiro	40°10'44.36"N	8°19'6.97"W
MC13	*Linaria incarnata*	PT: Coimbra, Santo António dos Olivais	40°13'35.71"N	8°24'23.24"W
SC01	*Linaria polygalifolia* subsp. *polygalifolia*	PT: Viana do Castelo, Carreço, Praia do Carreço	41º44'28.86"N	8º52'33.85"W
SC25	*Linaria polygalifolia* subsp. *polygalifolia*	PT: Cascais, Praia do Guincho	38º43'55.00"N	9º28'9.21"W
MC84	*Linaria polygalifolia* subsp. *polygalifolia*	PT: Figueira da Foz, Praia de Quiaios	40°12'33.74"N	8°53'47.98"W
COI14	*Linaria polygalifolia* subsp. *polygalifolia*	PT: Quiaios, Murtinheira	-	-
COI16	*Linaria saxatilis*	PT: Quiaios, Trevim	-	-
MNHN30	*Linaria spartea*	PT: Herdade da Defese, Montes Juntos (Capelins parish)	-	-
SC14	*Linaria spartea*	PT: SetúBAl, Pegões	38°41'1.51"N	8°37'13.88"W
JC03	*Linaria spartea*	PT: Leiria, Berlengas	39º24'54.95"N	9º30'23.18"W
JC07	*Linaria spartea*	PT: Leiria, Berlengas	39º24'57.28"N	9º30'24.62"W
MNHN32	*Linaria spartea*	PT: Road Torrão to Alfundão	-	-
MC44	*Linaria spartea*	PT: Aveiro, Sever do Vouga, Parada	40°46'44.24"N	8°17'22.49"W
MC89	*Linaria spartea*	PT: Figueira da Foz, Praia de Quiaios	40°12'56.79"N	8°53'19.09"W

MC88	*Linaria spartea*	PT: Figueira da Foz, Praia de Quiaios	40°12'51.65"N	8°53'9.04"W
JC18	*Linaria spartea*	PT: Montemor-o-Velho, Quinhendros, Vila Mota	40º10'22,36"N	8º42'27,31"W
MC101	*Linaria supina*	SP: Galiza, Folgoso de Caurel, Castillo de Carbedo	42°38'8.02"N	7° 7'27.07"W
MC29	*Linaria supina*	PT: Condeixa-a-Nova, Casmilo	40° 2'50.38"N	8°29'48.04"W
MC63	*Linaria supina*	PT: Serra d'Aires e Candeeiros, Casal das Pias	39°33'36.91"N	8°48'30.15"W
MC71	*Linaria thriornithophora*	PT: Arganil, Pomares	40°16'14.24"N	7°53'30.12"W
MC19	*Linaria thriornithophora*	PT: Coimbra, Agrelo	40°16'49.14"N	8°21'17.87"W
MC12	*Linaria thriornithophora*	PT: Aveiro, Sever do Vouga, Pessegueiro do Vouga	40°42'0.18"N	8°22'3.11"W
SC30	*Linaria thriornithophora*	SP: Galiza, Folgoso de Caurel, Santa Eufemia	42º34'12.67"N	7º11'49.68"W
MC99	*Linaria triornithophora*	SP: Galiza, Quiroga	42°32'3.45"N	7°13'28.51"W
MC103	*Melampyrum pratense* subsp. *latifolium*	SP: Galiza, Folgoso de Caurel, Moreda	42°36'45.96"N	7° 6'13.10"W
JC08	*Melampyrum pratense* subsp. *latifolium*	PT: Serra do Gerês, Portela do Leonte	41º46'0.63"N	8º8'37.12"W
MC56	*Misopates calycinum*	PT: Serra d'Aires e Candeeiros, Pragais	39°34'46.89"N	8°49'30.07"W
MC106	*Misopates orontium*	SP: Galiza, Folgoso de Caurel, Visuña	42°36'24.48"N	7° 3'59.48"W
MC16	*Misopates orontium*	PT: Coimbra, Caneiro	40°10'43.60"N	8°19'6.93"W
MC27	*Misopates orontium*	PT: Sever do Vouga, Pessegueiro do Vouga	40°42'33.53"N	8°21'36.29"W
MC74	*Misopates orontium*	PT: Coimbra, Assafarges, Palheira	40° 9'54.93"N	8°27'12.79"W
MC79	*Misopates orontium*	PT: Caldas da Rainha, Salir do Porto, miradouro	39°29'50.26"N	9° 9'21.47"W
SC13	*Mysopates orontium*	PT: SetúBAl, Pegões	38°41'1.51"N	8°37'13.88"W
SC26	*Mysopates orontium*	PT: Coimbra, Souselas	40º16'55.15"N	8º25'3.02"W
MC80	*Nothobartsia asperrina*	PT: Caldas da Rainha, Salir do Porto, miradouro	39°29'51.50"N	9° 9'18.83"W
MC104	*Odontite vernus*	SP: Galiza, Folgoso de Caurel, Visuña	42°36'24.58"N	7° 3'19.16"W
MC83	*Odontitella virgata*	PT: Figueira da Foz, Lagoa da Vela,	40°16'13.10"N	8°47'43.13"W
MC87	*Odontitella virgata*	PT: Figueira da Foz, Praia de Quiaios	40°13'20.52"N	8°51'53.51"W

MC100	*Parentucellia viscosa*	SP: Galiza, Folgoso de Caurel, Castillo de Carbedo	42°38'12.87"N	7° 7'29.23"W
MC66	*Parentucellia viscosa*	PT: Seia	40°25'45.99"N	7°42'38.94"W
MC48	*Parentucellia viscosa*	PT: Serra d'Aires e Candeeiros, Casal de Vale de Ventos	39°26'52.21"N	8°54'54.73"W
MC90	*Pedicularis sylvatica* subsp. *lusitanica*	PT: Serra do Gerês, Borrageiro	41°46'3.55"N	8° 8'49.06"W
MC45	*Pedicularis sylvatica* subsp. *lusitanica*	PT: Serra d'Aires e Candeeiros, Casal de Vale de Ventos	39°27'20.23"N	8°54'32.29"W
MC91	*Pedicularis sylvatica* subsp. *lusitanica*	PT: Serra do Gerês, Borrageiro	41º46'57.61"N	8º06'42.56"W
MC97	*Pedicularis sylvatica* subsp. *sylvatica*	PT: Serra do Gerês, Borrageiro	41º44'50.98"N	7º57'18.43"W
MC108	*Rhinanthus minor*	SP: Galiza, Folgoso de Caurel, Moreda	42°37'46.81"N	7° 6'34.18"W
MNHN34	*Scrophularia auriculata*	PT: Minas do Bugalho		
MC86	*Scrophularia frutescens*	PT: Figueira da Foz, Praia de Quiaios	40°12'35.70"N	8°53'46.17"W
MC81	*Scrophularia frutescens*	PT: Caldas da Rainha, Salir do Porto	39°30'4.78"N	9° 9'5.74"W
COI24	*Scrophularia frutescens*	PT: Quiaios, Murtinheira		
MC26	*Scrophularia grandiflora*	PT: Coimbra, Buçaco	40°22'37.10"N	8°21'57.66"W
MC18	*Scrophularia grandiflora*	PT: Coimbra, Vale de Canas	40°12'36.59"N	8°22'32.83"W
MC42	*Scrophularia grandiflora*	PT: Coimbra, Assafarges, Palheira	40° 9'57.75"N	8°27'5.48"W
MC50	*Scrophularia grandiflora*	PT: Serra d'Aires e Candeeiros, Casal de Vale de Ventos	39°25'15.09"N	8°55'44.65"W
SC20	*Scrophularia grandiflora*	PT: Coimbra, Jardim Botânico	40º12'24.69"N	8º25'19.27"W
UPT22	*Scrophularia grandiflora*	PT: Condeixa, Pitança de Baixo		
UPT23	*Scrophularia herminii*	PT: Cinfães, Ferreiros		
MC93	*Scrophularia lyrata*	PT: Serra do Gerês, Borrageiro	41º46'42.27"N	7º 56'17.70"W
MC58	*Scrophularia sambucifolia*	PT: Serra d'Aires e Candeeiros, Serro Ventoso	39°33'47.97"N	8°49'49.89"W
MC110	*Scrophularia scorodonia*	SP: Galiza, Folgoso de Caurel Seoane	42°38'20.47"N	7° 9'0.72"W
MC51	*Scrophularia scorodonia*	PT: Vale de Cambra, Souto Mau	40°46'53.07"N	8°16'50.46"W
MC22	*Scrophularia scorodonia*	PT: Coimbra, Buçaco	40°22'31.48"N	8°21'29.07"W

MC76	*Scrophularia scorodonia*	PT: Serra d'Aires e Candeeiros, Serro Ventoso	39°34'3.29"N	8°49'25.61"W
MC77	*Scrophularia scorodonia*	PT: Porto de mós, Lagoa de ArraBAl	39°29'54.36"N	8°52'18.75"W
MC46	*Scrophularia scorodonia*	PT: Serra d'Aires e Candeeiros, Casal de Vale de Ventos	39°27'4.36"N	8°54'40.90"W
UPT24	*Scrophularia scorodonia*	PT: Figueira da Foz, Santana		
JC02	*Scrophularia sublyrata*	PT: Leiria, Berlengas	39º24'54.99"N	9º30'23.18"W
JC06	*Scrophularia sublyrata*	PT: Leiria, Berlengas	39º24'45.21"N	9º30'40.86"W
MC07	*Verbascum levanticum*	PT: Coimbra, Jardim Botânico	40°12'23.42"N	8°25'16.02"W
MC85	*Verbascum litigiosum*	PT: Figueira da Foz, Praia de Quiaios	40°12'37.94"N	8°53'46.38"W
MC59	*Verbascum pulverulentum*	PT: Serra d'Aires e Candeeiros, Serro Ventoso	39°33'49.13"N	8°49'49.58"W
MC98	*Verbascum simplex*	SP: Galiza, Quiroga	42°32'3.75"N	7°13'25.17"W
MC68	*Verbascum simplex*	PT: Seia, Aldeia da Serra	40°24'51.73"N	7°41'28.11"W
MC36	*Verbascum simplex*	PT: Aveiro, Sever do Vouga, Parada	40°46'41.72"N	8°17'20.19"W
MC37	*Verbascum simplex*	PT: Aveiro, Sever do Vouga, Pessegueiro do Vouga	40°41'25.14"N	8°23'19.65"W
MC73	*Verbascum sinuatum*	PT: Coimbra, Carvalhais	40°10'53.94"N	8°26'35.26"W
MC31	*Verbascum sinuatum*	PT: Condeixa-a-Nova, Peixeiro	40° 4'36.17"N	8°30'19.46"W
MC57	*Verbascum sinuatum*	PT: Serra d'Aires e Candeeiros, caminho para a Fórnea	39°33'58.13"N	8°47'55.93"W
MC60	*Verbascum sinuatum*	PT: Serra d'Aires e Candeeiros, Casal das Pias	39°33'57.44"N	8°48'46.39"W
ISAUTL9	*Verbascum sinuatum*	PT: Lisboa	-	-
MNHN35	*Verbascum sinuatum*	PT: Road Juromelha to Elvas	-	-
MC60A	*Verbascum virgatum*	PT: Serra d'Aires e Candeeiros, Casal das Pias	39°33'57.44"N	8°48'46.39"W
MNHN36	*Verbascum virgatum*	PT: Road to Mourão	-	-
SC21	*Veronica acidifolia*	PT: Coimbra, Jardim Botânico	40º12'21.61"N	8º25'14.36"W
MC25	*Veronica arvensis*	PT: Coimbra, Buçaco	40°22'30.90"N	8°21'54.83"W
JL06	*Veronica arvensis*	PT: Coimbra, Jardim Botânico	40°12'23.40"N	8°25'27.68"W

MC53	*Veronica arvensis*	PT: Coimbra, Jardim Botânico	40°12'23.03"N	8°25'30.80"W
MC05	*Veronica arvensis*	PT: Coimbra, Jardim Botânico	40°12'23.11"N	8°25'30.85"W
MC09	*Veronica arvensis*	PT: Aveiro, Sever do Vouga, Parada	40°46'38.68"N	40°46'38.68"N
MC102	*Veronica chamaedrys* subsp. *chamaedrys*	SP: Galiza, Folgoso de Caurel, Moreda	42°36'51.14"N	7° 5'53.90"W
MC03	*Veronica hederifolia*	PT: Coimbra, Jardim Botânico	40°12'22.57"N	8°25'15.36"W
COI31	*Veronica micrantha*	PT: Gouveia	-	-
MC107	*Veronica officinalis*	SP: Galiza, Folgoso de Caurel, Moreda	42°36'40.89"N	7° 6'14.91"W
MC23	*Veronica officinalis*	PT: Coimbra, Buçaco	40°22'38.75"N	8°21'39.97"W
MC95	*Veronica officinalis*	PT: Serra do Gerês, Borrageiro	41º46'42.27"N	7º 56'17.70''W
JC05	*Veronica officinalis*	PT: Serra do Gerês, Borrageiro	41º45'43.12"N	8º7'50.40"W
JL07	*Veronica peregrina* subsp. *peregrina*	PT: Coimbra, Jardim Botânico	40°12'23.45"N	8°25'26.37"W
MC11	*Veronica persica*	PT: Aveiro, Sever do Vouga, Parada	40°46'38.48"N	8°17'22.64"W
MC39	*Veronica persica*	PT: Aveiro, Albergaria-a-Velha, Quinta da Lagoa	40°42'0.69"N	8°28'44.24"W
MC40	*Veronica persica*	PT: Aveiro, Sever do Vouga, Paradela	40°42'11.54"N	8°21'32.13"W
MC41	*Veronica persica*	PT: Coimbra, Assafarges, Palheira	40° 9'57.99"N	8°27'6.28"W
MC43	*Veronica persica*	PT: Coimbra, Condeixa-a-Nova	40° 5'8.85"N	8°28'3.51"W
MC55	*Veronica persica*	PT: Jardim Botanico	40°12'23.06"N	8°25'30.66"W
MC61	*Veronica persica*	PT: Serra d'Aires e Candeeiros, Casal das Pias	39°33'57.63"N	8°48'46.81"W
MC04	*Veronica polita*	PT: Coimbra, Jardim Botânico	40°12'23.12"N	8°25'30.67"W
MC64	*Veronica polita*	PT: Sever do Vouga	40°43'57.32"N	8°22'12.57"W

[1] The letters in the ID code refer either to the collector of the population (MC, Mariana Castro; SC, Sílvia Castro; JL, João Loureiro; JP, Jorge Paiva; JC, Joana Costa) or to the *index seminum* that sent us seeds (UPT, Universidade Portucalensis; COI, Universidade de Coimbra; ISAUTL, Instituto Superior de Agronomia, Universidade Técnica de Lisboa; MNHN, Museu Nacional de História Natural).

Appendix 2. Species, chromosomes numbers and distribution of Scrophulariaceae *s.l.* species occurring in Portugal.

Genus	Species	Subspecies	Chromosome numbers[1]	Distribution[2]
Anarrhinum	***duriminium***		18	Port: BA, DL, Mi, TM
	longipedicellatum		18	Port: BA BL DL
	bellidifolium		18	Port: AAl Ag BA BAl (BB) BL (DL) E Mi (R) TM
Antirrhinum	***lopesianum***		16	Port: TM
	meonanthum		16	Port: BA BB BL DL TM
	braun-blanquetii		16	Port: TM
	graniticum		16	Port: AAl Ag BA BAl BB BL TM
	onubense		16	Port: Ag
	linkianum		16	Port: BAl BB BL E R
	cirrhigerum		16	Port: Ag BAl E
Bacopa	*monnieri*		32/64/68	Port: [(BL)] [(Mi)]
Bartsia	*trixago*		24	Port: AAl Ag BA BAl BB BL (DL) E Mi R TM
Chaenorhinum	*origanifolium*	*origanifolium*	14	Port: BL E R TM
	segoviense	***segoviense***		Port: (TM)
	serpyllifolium	***lusitanicum***		Port: Ag
	minus		14	Port: BA (BAl) BB (BL) (DL) (TM)
Cymbalaria	*muralis*	*muralis*	14	Port: All provinces
Digitalis	***purpurea***	*purpurea*	48/56/112	Port: AAl Ag BA BAl BB BL DL E Mi R TM
		amandiana		Port: BA Dl TM
	thapsi		56	Port: (Ag) AAl BA BB BL (DL) (Mi) TM
	mariana	***heywoodii***	56	Port: AAl
Euphrasia	*hirtella*		22	Port: TM
	minima		44	Port: TM
Gratiola	*officinalis*		32	Port: AAl BA (BB) BL DL Mi TM
	linifolia		96	Port: AAl BA BAl BB BL DL E TM
Kickxia	*elatine*	*elatine*	36	Port: (BAl) BL (Mi)
		crinita	36	Port: DL
	spuria	*integrifolia*	18	Port: AAl Ag BL E
	lanifera		18	Port: AAl Ag E R
	cirrhosa		18	Port: AAl Ag BAl BB BL DL E R
	commutata	*commutata*	18	Port: (E)
Lathraea	*clandestina*		42	Port: (BA)
Limosella	*aquatica*		36/40	Port: DL
Linaria	***hirta***		12	Port: Ag AAl BAl
	supina	*supina*	12	Port: BL E DL?

		maritima		Port: DL? Mi?
	polygalifolia	***polygalifolia***	12	Port: BL DL E Mi
		lamarckii	12	Port: Ag BAl E
	aeruginea	***aeruginea***	12	Port: BA E TM
	oblongifolia	***haenseleri***	n=6	Port: Ag AAl BAl
	ricardoi			Port: AAl Bal
	amethystea	***amethystea***	12	Port: AAl Ag BA BAl BL TM
		multipunctata		Port: BL E
	munbyana			Port: Ag
	simplex		12	Port: (BL)
	micrantha		12	Port: AAl (Ag) (BAl) (TM)
	saxatilis		12	Port: (AAl) BA BB BL DL E Mi R TM
Linaria	***bipunctata***	***bipunctata***		Port: DL TM
		glutinosa		Port: (Ag) BAl E R
	intricata			Port: DL TM
	diffusa		12	Port: AAl BA BB BL E TM
	triornithophora		12	Port: AAl BA BB BL DL Mi TM
	incarnata		12	Port: AAl (Ag) BB BL
	spartea		12	Port: AAl Ag BA BAl BB BL DL E Mi R TM
	algarviana		12	Port: Al
	viscosa	***viscosa***	12	Port: Ag BAl BL E
	pedunculata		12	Port: Ag BAl
	elegans		12	Port: BA BB BL Mi TM
Lindernia	*procumbens*		30	Port: Mi
	dubia		18/20/32	Port: [AAl] [BAl] [BL] [Mi] [R]
Melampyrum	*pratense*	*latifolium*	18	Port: (BA) (DL) Mi TM
Mimulus	*moschatus*		32	Port: [(R)]
Misopates	*orontium*		16	Not reported for Portugal in Flora Iberica
	calycinum			Port: Ag BAl E
Nothobartsia	*asperrima*		36	Port: Ag BAl BL DL E Mi R
Odontitella	***virgata***		26	Port: A Al Ag BA BAl BL DL E R TM
Odontites	*viscosus*	*australis*	20/22/24	Port: E R
	vernus		18/20/38/39/40	Port: BA BB DL Mi TM
Parentucellia	*viscosa*		48	Port: AAl Ag BA BB BL (DL) E Mi R TM
	latifolia		48	Port: AAl Ag BA BB (BL) E Mi R TM
Pedicularis	*sylvatica*	*lusitanica*	16	Port: AAl BA BAl BB BL DL E Mi E TM
		sylvativa	16	Port: TM*
Rhinanthus	*minor*		14/22	Port: BA Mi TM
	nodosa		36	Not reported for Portugal in

Scrophularia				Flora Iberica
	bourgaeana		42	Port: Mi
	herminii		52/68	Port: BA (DL) Mi TM
	lyrata		58	Port: AAI Ag BB E R TM
	auriculata	*auriculata*	78/80/84/86/88	Port: AAI Ag BA BAI BB BL DL E Mi R TM
Scrophularia	*scorodonia*		58/60-80	Port: AAI Ag BA BAI BB BL (DL) E Mi R TM
	sublyrata		58/60	Port: AAI Ag BA BAI BB (DL) E Mi TM
	sambucifolia	*sambucifolia*	52/58	Port: AAI Ag BAI E R
	grandiflora		58	Port: BL
	valdesii		58	Port: TM
	peregrina		36	Port: BB E R
	canina	*canina*	24/26/30	Port: AAI Ag BA BAI BB BL (DL) E Mi R TM
	frutescens		26	Port: Ag BAI BB BL DL E Mi
Sibthorpia	*europaea*		18	Port: AAI Ag BA (BAI) BB BL DL E Mi TM
	perefrina		20	Port: [E]
Verbascum	***barnadesii***		48/55-58/62	Port: AAI Ag (BAI) BB E R
	levanticum		44/48	Port: [(BA)] [BL]
	virgatum		62/64/66	Port: AAI Ag BA BAI BB BL DL E Mi R TM
	simplex		32	Port: AAI Ag BA BAI BB BL DL E Mi TM
	thapsus		32/34/36	Port: TM
	giganteum	*martinezii*		Port: BAI
	litigiosum		36	Port: Ag BAI BL (DL)
	pulverulentum		32	Port: AAI BA BB BL R TM
	sinuatum		18/24/30	Port: AAI Ag (BA) (BAI) BB BL (DL) E Mi R TM
Veronica	*serpyllifolia*		14/16/28	Port: AAI BA BL DL Mi TM
	nevadensis		14	Port: BA
	officinalis		18/36	Port: BA (BB) BL DL Mi TM
	scutellata		18	Port: BA BB BL DL E Mi TM
	montana		18/36	Port: BA BL Mi
	chamaedrys	*chamaedrys*	16/32	Port: DI (Mi) TM
	micrantha		16	Port: BA BB BL (DL) Mi R TM
	arvensis		16	Port: AAI Ag BA (BAI) BB BL DL E Mi R TM
	verna		16	Port: BA
	triphyllos		24	Port: BA BL TM
	polita		14	Port: AAI Ag (BB) BL E R TM
	agrestis		28	Port: BA BL (E) (R)
	persica		28	Port: AAI (Ag) (BAI) Ba BL DI E (Mi) R TM
	cymbalaria		36/54	Port: AAI BAI (BB) E R
	hederifolia		36/54/56	Port: AAI Ag BA BAI BB BL

				E R TM
	acinifolia		14/16	Port: AAl (Ag) BA BL DL (E) R TM
Veronica	*peregrina*	*peregrina*	52	Port: [BL] [(DL)] [E] [(Mi)] [R]
	beccabunga	*beccabunga*	16/28/36	Port: BA BB BL (Mi) TM
	anagallis-aquatica	*anagallis-aquatica*	34/35/36	Port: AAl Ag BA BAl BB BL DL E R TM
	anagalloides	*anagalloides*	18/18+2B/36	Port: BAl BL E TM

* First documentation in Portugal

[1] Chromosome numbers according to Flora Iberica (Benedí et al. 2009), to Tropicos® (http://www.tropicos.org/) and Anthos (http://www.anthos.es/) online databases and to M. Queirós printed files (available at the Department of Life Sciences, FCTUC)

[2] Distribution among Portuguese provinces according with Flora Iberica (Benedí et al. 2009). Provinces in round brackets indicate that the information on the occurrence of a given *taxon* in that province was provided by a specialist and not from material analysed by the author of the revision. Provinces in square brackets indicate that the *taxon* is naturalized in that province. Provinces followed by a question mark indicate that the presence of the *taxon* in that province is uncertain. Province abbreviations: AAl, Alto Alentejo; Ag, Algarve; BA, Beira Alta; BAl, Baixo Alentejo; BB, Beira Baixa; BL, Beira Litoral; DL, Douro Litoral; E, Estremadura; Mi, Minho; R, Ribatejo; TM, Trás-os-Montes e Alto Douro.

Appendix 3. Statistical analyses on genome size estimates among the sampled species of each genera and among species from closely related genera.

Taxa	n	*δ*	Statistically test	*P*
Anarrhinum spp.	52	2	*F* = 1.51	0.320
Antirrhinum spp.	27	4	*F* = 2.39	0.081
Bartsia sp., *Nothobartsia* sp., *Parentucellia* sp.	24	2	*F* = 300.76	< 0.001
Digitalis spp.	33	3	*F* = 129.93	< 0.001
Linaria spp.	72	8	*F* = 750.99	< 0.001
Misopates spp.	22	-	*t* = 0.01	0.991
Odontite sp., *Odontitella* sp.	9	-	*t* = 206.23	< 0.001
Pedicularis spp.	10	-	*t* = 3.01	0.017
Scrophularia spp.	77	8	*H* = 62.72	< 0.001
Verbascum spp.	39	5	*F* = 374.31	< 0.001
Veronica spp.	54	7	*F* = 1677.4	< 0.001

n represents the number of samples and *δ* the degrees of freedom. A t-test was applied to *Misopates* spp., *Odontite* sp. *vs. Odontitella* sp. and *Pedicularis* spp.; a non-parametric Kruskal-Wallis one-way ANOVA on ranks was applied to *Scrophularia* spp., while a one-way ANOVA was applied to the remaining *taxa* (for details on statistical analyses see *Materials and Methods*).