UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Agronomia Animali Alimenti Risorse Naturali e Ambiente

---

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE DELLE PRODUZIONI VEGETALI

CICLO: XXVIII

BREEDING F1 HYBRID VARIETIES OF LEAF CHICORY THROUGH MARKER-ASSISTED SELECTION SCHEMES

Direttore della Scuola : Ch.mo Prof. Antonio Berti
Supervisore :Ch.mo Prof. Gianni Barcaccia

Dottorando : Andrea Ghedina

Declaration

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Andrea Ghedina, 18/07/2016

A copy of the thesis will be available at *http://paduaresearch.cab.unipd.it/*

Dichiarazione

Con la presente affermo che questa tesi è frutto del mio lavoro e che, per quanto io ne sia a conoscenza, non contiene materiale precedentemente pubblicato o scritto da un'altra persona né materiale che è stato utilizzato per l'ottenimento di qualunque altro titolo o diploma dell'università o altro istituto di apprendimento, a eccezione del caso in cui ciò venga riconosciuto nel testo.

Andrea Ghedina, 18/07/2016

Una copia della tesi sarà disponibile presso *http://paduaresearch.cab.unipd.it/*

Index

Riassunto

Il genere *Cichorium* (*Cichorium intybus* subsp. *intybus* var. *foliosum* L.) comprende specie vegetali diploidi (2*n*=18) appartenenti alla famiglia delle Asteraceae.

Queste specie sono generalmente biennali o, in natura, perenni. Si tratta di specie prevalentemente allogame a causa di un efficiente sistema di autoincompatibilità sporofitica. Inoltre, la fecondazione incrociata è favorita da una morfologia fiorale sfavorevole all'autofecondazione, se in assenza di un impollinante (ad esempio la proterandria, ovvero la presenza di antere mature prima dei pistilli). In specie di questo genere si riscontra, inoltre, una forte competizione gametofitica, ovvero il rigetto di polline geneticamente simile a quello prodotto dal parentale femminile.

A lungo apprezzate come piante medicinali dagli antichi Greci e Romani, le specie del genere *Cichorium* sono attualmente tra le più importanti specie orticole coltivate. Sono generalmente utilizzate nella preparazione di insalate fresche o, più raramente, cucinate secondo le tradizioni alimentari locali.

Anche se la coltivazione della cicoria non contribuisce largamente all'introito agricolo totale di ogni paese, questa risulta essere molto importante a livello locale, in quanto caratterizza l'agricoltura di aree limitate dove è concentrata la coltivazione dell'80-90% della produzione nazionale di questa coltura. Questo è in effetti il caso dell'Italia dove, il 66% dell'ettarato e il 59% della produzione nazionale di una cicoria da foglia rossa o variegata conosciuta come "Radicchio" è concentrata in Veneto.

La produzione di Radicchio è stata per un lungo periodo basata su popolazioni mantenute dagli agricoltori, che venivano di anno in anno selezionate e il cui seme veniva solitamente riutilizzato all'interno del centro aziendale o poteva venir venduto attraverso transazioni private e solitamente non ufficiali.

Tutte queste popolazioni, ottenute per selezione massale e mantenute attraverso l'interincrocio di parentali selezionati, devono essere considerate altamente eterozigoti e geneticamente eterogenee; il comportamento e il livello di adattamento di queste popolazioni ai diversi ambienti e condizioni agronomiche, è funzione della frequenza di geni favorevoli e della loro combinazione.

In ogni programma di miglioramento genetico, gli schemi e i metodi di selezione che possono essere adottati e le tipologie varietali che possono venir costituite, dipendono dalle

barriere riproduttive delle piante (come l'autoincompatibilità), dai sistemi di impollinazione (ad esempio l'allogamia) così come dalla struttura genetica delle popolazioni.

La forte autoincompatibilità rinvenuta in cicoria, rende estremamente difficile l'ottenimento di parentali altamente omozigoti nell'ottica di un efficiente schema di ottenimento di ibridi F1.

Nonostante le difficoltà incontrate nell'ottenimento di linee inbred attraverso cicli ripetuti di autofecondazione, la recente scoperta di mutanti maschiosterili spontanei ha aumentato l'interesse verso la produzione di varietà ibride F1.

La maschiosterilità, o l'incapacità delle piante di produrre polline funzionale, è utile per la produzione di seme ibrido commerciale, attraverso l'incrocio di linee inbred parentali appropriatamente selezionate attraverso test di progenie, per stabilire l'attitudine alla combinazione specifica.

In questo progetto abbiamo sviluppato un metodo di genotipizzazione utilizzando marcatori molecolari, utile per stabilire il grado di omozigosi e la stabilità genetica di singole linee inbred e per la misurazione dell'attitudine alla combinazione specifica tra linee portaseme ed impollinante sulla base della loro diversità genetica. Queste informazioni possono essere utilizzate per programmare incroci e predire l'eterosi di ibridi F1 sperimentali sulla base della distanza genetica della linee parentali. Conoscendo il genotipo dei parentali, saremo in grado non solo di proteggere ogni nuova varietà rilasciata sul mercato ma anche di stabilire il grado di purezza e identità varietale del seme ibrido commerciale e di certificare l'origine dei suoi derivati alimentari.

La costituzione di varietà ibride F1 vigorose e stabili si avvarrà del breeding moderno assistito da marcatori utilizzando tecnologie quali marcatori SSR e SNP, senza alcun collegamento con tecniche di l'ingegneria genetica, combinando tratti di qualità, uniformità e produttività all'interno degli stessi genotipi.

Inoltre, nell'ambito di questo progetto di ricerca abbiamo trattato la scoperta e la caratterizzazione genetica di quattro mutanti maschiosterili in questa specie. Questi mutanti, che da quanto ci risulta dalle informazioni in nostro possesso sono i primi mutanti maschiosterili spontanei mai scoperti e descritti in Radicchio, sono stati dettagliatamente caratterizzati per la via metabolica dello sviluppo della microsporogenesi e gametogenesi e per il modello di ereditarietà del gene responsabile della maschiosterilità. Abbiamo

sviluppato un saggio diagnostico basato su marcatori molecolari per la selezione precoce di genotipi associati a piante maschiosterili. In questo modo si è dimostrato che i mutanti maschiosterili oggetto di questo progetto di ricerca sono controllati da un singolo gene nucleare (*ms1*) che agisce allo stato recessivo. Siamo stati in grado di mappare il gene responsabile della maschiosterilità in una regione cromosomica satura e ben caratterizzata di circa 7.3 cM e 5.8 cM dal locus *ms1*.

Complessivamente queste informazioni saranno fondamentali per la pianificazione di un esperimento di Genotyping-by-Sequencing basato su una popolazione BC1 con l'obiettivo di restringere la finestra genomica contenente il gene responsabile della maschiosterilità in cicoria da foglia.

Infine il sequenziamento e l'assemblaggio della prima bozza del genoma di cicoria da foglia discussi in questo studio, contribuiranno ad aumentare e rafforzare la credibilità delle ditte sementiere italiane e delle attività locali in Veneto correlate alla coltivazione e la commercializzazione del Radicchio e prodotti alimentari connessi; il mercato sementiero di questa specie avrà la possibilità di diventare altamente professionale e fortemente competitivo a livello nazionale ed internazionale.

Abbiamo assemblato una bozza del genoma della misura stimata di 760 Mb. Abbiamo ottenuto 58.392.530 e 389.385.400 sequenze grezze rispettivamente attraverso le piattaforme di sequenziamento MySeq e HiSeq. Abbiamo identificato complessivamente 66.785 regioni contenenti motivi SSR. Abbiamo riportato i dati bioinformatici di assemblaggio della prima bozza del genoma di Radicchio, insieme ai dati più salienti provenienti da una predizione genica *de novo* e da una annotazione funzionale *in silico* di più di 18.000 unigenes.

La scoperta del genoma di Radicchio significa l'ottenimento di una base scientifica solida e una conoscenza tecnologica tale da giocare in breve tempo un ruolo cruciale nella risoluzione di problematiche legate alla protezione e la coltivazione di moderne varietà di Radicchio.

Siamo fiduciosi che i nostri sforzi amplieranno l'attuale conoscenza dell'organizzazione del genoma e della composizione genica della cicoria da foglia, che si ritengono essere fondamentali nello sviluppo di nuovi strumenti e saggi diagnostici basati su marcatori

molecolari utili alle strategie di costituzione varietale e che permettano studi più specifici di tratti cromosomici che controllino aspetti agronomici rilevanti in questa specie.
In conclusione, il presente lavoro è da leggersi come una sorta di manuale per meglio capire il mondo di una specie non modello come la cicoria da foglia, orientato prevalentemente verso i costitutori genetici e i produttori di sementi di cicoria.

Summary

Cichorium (*Cichorium intybus* subsp. *intybus* var. *foliosum* L.) comprises diploid plant species (2*n*=18) belonging to the Asteraceae family.

These species are biennial or, in the wild, perennial species. They are naturally allogamous due to an efficient sporophytic self-incompatibility system. In addition, outcrossing is promoted by a floral morpho-phenology unfavorable to selfing in the absence of pollen donors (*i.e.*, proterandry, wherein the anthers mature before the pistils) and a favorable competition of allo-pollen grains and tubes (*i.e.*, pollen that is genetically diverse from that produced by the seed parents, usually called auto-pollen).

Long appreciated as medical plants by the ancient Greeks and Romans, *Cichorium* spp. are currently among the most important cultivated vegetable crops. They are generally used as components in fresh salads or, more rarely, cooked according to local traditions and alimentary habits.

Although this crop does not contribute greatly to the total agricultural income of each country, it is very important at the local level, as it characterizes the agriculture of limited areas, where from 80 to 90% of the country's production is concentrated. This is indeed the case of Italy, where the Veneto region accounts for 66% of the national acreage and 59% of the national production of the particular type of red or variegated chicory known as "Radicchio".

Radicchio production was for a long time based on farmer's populations, which are yearly selected and maintained and whose seed is usually reutilized on farm but may also be sold through private and not officially registered transactions. All these populations, obtained by mass selection and maintained through the inter-crossing of selected parents, have to be considered highly heterozygous and genetically heterogeneous whose behaviour and level of adaptation to different environments and/or cultural conditions depend on the frequency of favourable genes or gene combinations.

In each breeding program, selection schemes and methods that can be used and the varietal types than can be breed, depend on plant reproductive barriers (*e.g.* self-incompatibility) and pollination system (*e.g.* allogamous), and thus on the genetic structure of populations.

As a matter of fact, the strong self-incompatibility system found in chicory hinders obtaining highly homozygous parents, made it generally difficult to propose an efficient F1 seed production scheme.

Despite the difficulties encountered in obtaining inbred lines by repeated selfing, the recent discovery of spontaneous male-sterile mutants increased the interest towards the production of F1 hybrid varieties. Indeed, male-sterility, or the inability of plants to produce functional pollen, is needful to the commercial production of hybrid seed by crossing parental inbred lines appropriately selected through progeny tests for assessing their specific combining ability.

In this project we developed a genotyping method using molecular markers, useful for assessing the homozygosity and genetic stability of single inbred lines and for measuring the specific combining ability between maternal and paternal inbred lines on the basis of their genetic diversity. This information could be exploited for planning crosses and predicting the heterosis of experimental F1 hybrids on the basis of the allelic divergence and genetic distance of the parental lines. Knowing the parental genotypes would enable not only to protect newly registered varieties but also to assess the genetic purity and identity of the seed stocks of commercial F1 hybrids, and to certificate the origin of their food derivatives.

Modern marker-assisted breeding (MAB) technology based on traditional methods using molecular markers such as SSRs and SNPs, without relations to genetic modification (GM) techniques, will now be planned and adopted for breeding of vigorous and uniform F1 hybrids combining quality, uniformity, and productivity traits in the same genotypes.

Furthermore, this research project deals with the discovery and genetic analysis of four male-sterile mutants in this species. These mutants, which to the best of our knowledge are the first spontaneous male-sterile mutants ever discovered and described in Radicchio, were characterized in great details for the developmental pathway of micro-sporogenesis and gametogenesis, and the inheritance pattern of the gene underlying the male-sterility trait. A quick molecular diagnostic assay was also developed for the early marker-assisted selection of the genotype associated to male-sterile plants. Hence, male-sterile mutants object of this PhD project were demonstrated to be controlled by a single nuclear gene (*ms1*) that acts at the recessive status. We were able to map the male-sterility gene on a well saturated and

characterized linkage group in a chromosomal region spanning 7.3 cM and 5.8 cM from the *ms1* locus. On the whole, this information was crucial to plan a Genotyping-by-Sequencing experiment based on BC1 progenies with the aim of narrowing down the genomic window containing the gene for male-sterility in leaf chicory.

Finally, the sequencing and assembly of the first genome draft of leaf chicory, will contribute to increase and reinforce the reliability of Italian seed firms and local activities of the Veneto region associated with the cultivation and commercialization of Radicchio plant varieties and food products; the seed market of this species will have the chance to become highly professional and more competitive at the national and international levels.

We assembled a genome draft of an estimated size of 760 Mb. We obtained 58,392,530 and 389,385,400 raw reads through the MySeq and HiSeq platforms, respectively. Overall, we identified 66,785 SSR containing regions. Original data from the bioinformatic assembly of the first genome draft of Radicchio, along with the most relevant findings that emerged from an extensive *de novo* gene prediction and *in silico* functional annotation of more than 18,000 unigenes are critically discussed.

To uncover the sequence of a given genome means to gain a robust scientific background and technological knowhow, which in short time can play a crucial role in addressing and solving issues related to the cultivation and protection of modern Radicchio varieties. In fact, we are confident that our efforts will extend the current knowledge of the genome organization and gene composition of leaf chicories, which is crucial in the development of new tools and diagnostic markers useful for our breeding strategies, and allow researchers for more focused studies on chromosome regions controlling relevant agronomic traits of Radicchio.

In conclusion, the present work is a sort of handbook to better understand the world of a non-model species, *i.e.* leaf chicory, and it is mainly directed to breeders and seed producers dealing with leaf chicory.

Chapter I

Current advances in genomics and breeding of leaf chicory (*Cichorium intybus* L.)

Submitted: Barcaccia G, Ghedina A, Lucchin M, Current advances in genomics and breeding of leaf chicory (*Cichorium intybus* L.), Agriculture.

Abstract

This review gives an overview of agricultural topics on a non-model species, *i.e.* leaf chicory. Often classified as a minor crop, “Radicchio”, the Italian name of leaf chicory, is assuming a very important role at both local and national level, as it characterizes a high proportion of the agricultural income of suited areas. Botanical classification along the genus *Chichorium* is reported and a detailed description of the most important cultivated biotypes typical of Northern Italy is presented. A special consideration is reserved to breeding aspects, from molecular marker-assisted selection to the implementation of the first genome draft and leaf transcriptomes. Sexual barriers, *e.g*. self-incompatibility or male-sterility, are described in great details with the aim to be utilized for breeding purposes. The main aspects of seed production and aspects are also critically presented. In conclusion, the present work is a sort of handbook to better understand this orphan crop and it is mainly directed to breeders and seed producers dealing with leaf chicory.

Keywords: Radicchio, population genetics, cultivated varieties, marker-assisted breeding.

Introduction

A leafy vegetable, chicory is among the most well-known and popular horticultural plants in the world. Although there are large differences with regard to cultural practices and utilizations, chicory is found in almost every country and is included in the diet of most Western as well as Eastern populations.

Chicory is a traditional European horticultural crop, and although it cannot be considered autochthonous, its evolution as a vegetable crop occurred on continental Europe, where it gradually differentiated into a variety of cultivated types. According to Street *et al.* [1], *Cichorium intybus* can be divided in four varieties according to the purpose and use for which it was cultivated: i) "industrial" or "root" chicory, which is predominantly cultivated in north-western Europe, India, South Africa and Chile. At present, the taproot utilization of this type appears to be mainly for inulin extraction or, on a more limited scale, for the production of a coffee substitute; ii) "forage" chicory, a variety that was initially derived from wild chicory, has been domesticated since the mid-1970s to intensify herbage obtainability in perennial pastures for livestock; iii) "brussels" or "Witloof" chicory, which is commonly grown in Europe as an industrial crop for etiolated buds obtained by forcing; iv) "leaf" chicory, also called "Radicchio", which is mainly known as an important component of fresh salads and is often cooked and differently prepared according to traditions and alimentary habits [2].

Currently, *C. intybus* is mainly grown throughout continental Europe, in South Western Asia, and in limited areas of Northern America, South Africa, and Australia. Most likely known by the Egyptians as a medicinal plant [3] and used as a vegetable crop by ancient Greeks and the Romans [4], chicory gradually underwent a process of naturalization in Europe. Although it cannot be considered as an autochthonous species, it became part of the natural and agricultural European flora. Wild *C. intybus* covers a great portion of the entire European continent; it has traditionally become a part of the diet of local populations as an important ingredient of typical local dishes. This might be both the consequence and the cause of the great differentiation among a number of types, which have originated an ever-increasing number of cultivar groups, types, and populations that altogether comprise the horticultural landscape of the genus *Cichorium*. This genus is particularly rich and interesting from historical, cultural, agronomical, commercial, and scientific points of view.

The commonly named “root chicory” appears to derive from so-called “Magdeburg chicory”, the ancient root chicory known and traditionally used in some European countries as a coffee substitute since the end of the 16th century and that surged to outstanding importance with the continental block at the time of Napoleon. Also a very important leafy vegetable, so-called “Witloof Chicory” or “Belgian Endive”, perhaps the best known among leafy chicories, is considered to be a derivative of Magdebourg chicory. It is commonly accepted that the first well-known pale-yellowish sprouts of Witloof chicory were casually obtained by a Belgian farmer around 1870s, who had observed and harvested the sprouts from a stock of roots piled up in autumn and left apart during the cold season, waiting to be dried, grounded, and toasted [2].

Lacking comprehensive, homogeneous, sufficiently detailed, and univocal data regarding horticultural production and trade, it is impossible to provide reliable figures on the diffusion and economic importance of the crop in Europe, where chicory is mostly grown [2]. The most recent statistics concerning the European market [5] often include chicory under the general declaration “salads”, or it is considered together with lettuce, which is by far the most important leafy vegetable at both European and world-wide scales. The situation is not very different if one considers, as a source of reliable information, the statistics of each single country. Nonetheless, on the basis of accessible data, it is possible to determine that Italy is the almost exclusive producer of chicory. Although this crop does not contribute greatly to the total agricultural income of each country, it is very important at the local level, as it characterizes the agriculture of limited areas, where from 80 to 90% of the country’s production is concentrated. This is indeed the case for Italy, where the Veneto region accounts for 66% of the national acreage and 59% of the national production of the particular type of red or variegated chicory known as “Radicchio”. The cultivation of other chicory types such as chicory of Catalogne in both France and Italy is much less concentrated and may extend far south. Furthermore, it is worth noting that chicory is not only important for the local economy but it may also have significance at an international trade scale. Altogether, the US imports of chicory in 2011 [5] equaled 174,970,000 Kg, with a value of $ 185,131,000. Approximately half of the amounts, both quantity and value, are represented by Witloof chicory, for which import from Belgium and The Netherlands reached more than 90% of the total. Thus, although still grown at a regional scale, Chicory

has a place among more known and used vegetables and may represent a non-negligible source of income for farmers in areas where it has traditionally been grown.

Within this framework, two observations can be made. The first concerns the marked decrease in US not-qualified chicory imports from Europe, in particular from Belgium, The Netherlands and Italy (2,622 Mt in 1996, 536 Mt in 2002, 1,856,000 Kg in 2011), which corresponds with the increase in imports from Central and Southern America (1,046 Mt in 1996, 2,522 Mt in 2002, 149,310,000 Kg in 2011). Compared to the stable or slightly increasing figures recorded during the same period for Witloof chicory (between 2,000 and 2,400 Mt), this trend appears to indicate that Witloof has benefited from the quality and standardization of the marketable product. The second observation regards Radicchio, which is now considered with increasing attention both in Europe and in the US, as well as abroad, where its cultivation started some years ago and appears to have an improving evaluation, whereby the red or variegated leaves are particularly appreciated as a component of ready-to-eat salads.

In Italy, Radicchio of Chioggia is cultivated on a total area of approximately 16-18,000 ha, half of which is in the Veneto region, with a total production of approximately 270,000 tons (more than 60% obtained using professional seeds), reaching an overall turnover of approximately € 10,000,000 per year [6].

Taxonomy and Biology of Chicory

Chicory (*C. intybus* L.) belongs to the family Asteraceae, a very large family with approximately 23,000 species subdivided into 1,535 genera grouped into three subfamilies: Barnadesioideae, Cichorioideae and Asteroideae [7].

The tribe Lactuceae, in the subfamily Cichorioideae, includes the genus *Cichorium*, within which different species are recognized according to the origin. Referring to European flora, Tutin *et al.* [8] describe the three species *C. spinosum*, *C. intybus*, and *C. endivia* and subdivide the last into subsp. *endivia* (cultivated) and subsp. *divaricatum* (wild). Pignatti [9], taking into account Italian flora, refers to the three wild species as *C. spinosum*, *C. intybus*, with the var. *glabratum* (Presl) Fiori, and *C. pumilum*, considering *C. endivia* only as a cultivated species.

Integrating morphological characters with molecular observations, Kiers [10] describes the two cultivated and most known species, *C. intybus* and *C. endivia*, and the two wild species, *C. spinosum* and *C. pumilum*. Moreover, two more species, never observed in Europe, are added, *C. calvum* and *C. bottae*: the first is endemic to the dry and hot environments of the Middle East and South Western Asia and the second to Yemen and Saudi Arabia. More recently, Conti *et al.* [11] recognized three species in the genus with regard to Italian flora: *C. endivia*, with the two subspecies *endivia* Hegi and *pumilum* (Jacq) Cout.; *C. intybus*, with the two subspecies *glabratum* (C. Presl) Arcang. and *intybus*; and C. *spinosum*.

Since the early 1990s, when the analysis of DNA polymorphisms became more familiar to taxonomists, several studies attempted to explore, and possibly clarify, the relationships among the cultivated species - *C. intybus* - and its wild relative. Using mitochondrial RFLP markers, Vermeulen *et al*. [12] suggested that *C. spinosum* could be considered an ecotype of *C. intybus* rather than a separate species. Applying other and more informative molecular methods, such as nuclear ITS and SSR markers, Gemeinholzer and Bachmann [13] were unable to discriminate between these two species, which nonetheless could be clearly delimited with two diagnostic and one overlapping morphological character. On the basis of chloroplast DNA and nuclear rDNA sequence analysis [14] or AFLP fingerprints [15], it was confirmed that *C. intybus* is closely related to *C. spinosum*, whereas *C. endivia*, *C. pumilum* and *C. calvum* revealed a high genetic similarity with each other and are fairly

well separated from *C. intybus* and *C. spinosum*. The sixth species, *C. bottae*, must be considered a sister species.

In addition to morphological descriptors and molecular similarity or diversity estimates, a distinction among the above six *Cichorium* species can be made on the basis of their life cycle and reproductive system. Thus, two groups may be established: on one side, *C. intybus*, *C. spinosum* and *C. bottae*, which are perennials and characterized by a strong self-incompatibility system, on the other, *C. endivia*, *C. pumilum* and *C. calvum*, which are annual species and show self-compatibility. Although the names of the recognized botanical varieties do not appear within this framework, the various cultivated types have originated from these species.

The origin and differentiation of the genus is concordantly located in South-Eastern Europe, the eastern Mediterranean basin and South-west Asia. Within this large region, the area of origin of *C. intybus* tends to be mapped to the southern Balkan Peninsula and the northern Middle East. From there, it firstly migrated throughout the entire Mediterranean basin and toward southern and eastern Asia, where it appears to have found different areas of diffusion and adoption as horticultural crops.

Variously extensive lists of *Cichorium* species, subspecies, botanical varieties, and cultivar groups have been published; such lists are present on accessible internet sites, where scientific and technical information is often confused with commercial promotion, forming a mass of information that is not always easily interpreted. The most complete list appears to be the GRIN (Germplasm Resources Information Network) database released by the USDA [16], in which 212 entries of the genus *Cichorium* and 127 of the species *C. intybus* are cited. A brief description is given; in most cases, pictures of a not-well grown plant are present plus an actual short, though not always correct, botanical description (*i.e.*, genus, species and family). Many of names are synonyms and often refer to differences among commercial types rather than to taxonomic distinctiveness. A smaller list is provided in Mansfeld's World Database of Agricultural and Horticultural Crops [17], which includes 49 entries but only 15 if we consider just the species *C. intybus*.

On the basis of Kiers' [10] findings, Raulier *et al*. [18] presented results of genetic diversity within *C. endivia* and *C. intybus* using SSR markers. The study showed that the two species are strongly differentiated, even though some *C. intybus* individuals were genetically closer

to *C. endivia,* revealing complex genetic relationships between the two species. Regarding *C. intybus,* the results confirmed its differentiation into three cultivar groups (*i.e.,* Witloof, root chicory and leaf chicory). The sub-classification of leaf chicory into Radicchio, Sugarloaf and Catalogne previously based on morphological factors was also confirmed. Altogether, at least six cultivar groups, mainly differentiated on the basis of their use, are recognizable [14-15-18-19]. A summary is proposed in **Table 1**, where correspondence among taxonomy, cultivar group and most frequent and known utilization is attempted.

Table 1. Taxonomy of *Cichorium intybus*, including subspecies, botanical varieties and cultivar groups (adapted from Lucchin *et al*. [2]).

Taxonomy	Cultivar group	Utilization
C. intybus		
subsp. *intybus*	Wild	
var. *foliosum*	Witloof chicory	cooked/salads
var. *porphyreum*	Pain de Sucre	cooked/salads
var. *latifolium*	Radicchio	salads
var. *sylvestre*	Catalogne	cooked
var. *sativum*	Root chicory	coffee substitute (roasted) inulin extracts/cooked
subsp. *glabratum*	Wild	

Although chicory is cultivated throughout Europe and tends to expand toward new horticultural areas, most of the cultivar groups are well known. These groups have been extensively adopted as horticultural crops at a local scale only, and as such, their description in the scientific and technical literature is partially complete or not entirely accurate. This is particularly true when the crop is largely differentiated, with subgroups among which it may be very difficult to identify differences and affinities. We are aware that diversity, particularly for this type of horticultural crop, is often the most efficient tool for commercial success, both for the seed producers and the farmers. Thus, in a continuously changing breeding world, it may be unwise to establish a rigid framework within which everything must be ordered. Nevertheless, it seems advisable to have an account, as complete as possible, of the plant materials of chicory.

Five main groups can be recognized within *C. intybus* subspecies *intybus,* to which all the cultivated types of chicory belong, as stated by Kiers [10] and re-marked by Raulier *et al*.

[18]. Excepting wild accessions, the first referring to the var. *foliosum*, traditionally includes Witloof chicory. It might be argued that if Witloof chicory was obtained from the roots of Magdebourg, as commonly accepted, then it should be grouped under var. *sativum*, together with all the other root types. Nevertheless, most scientific literature refers to Witloof chicory as a type belonging to var. *foliosum* [2-10-18-19-20-21-22-23]. The cultivar groups Pain de Sucre, Radicchio and Catalogne are assigned to var. *porphyreum*, var. *latifolium* and var. *sylvestre*, respectively.

Italian Radicchio Biotypes and Varieties

Here we only discuss Radicchio (*i.e.,* var. *latifolium*). The Italian common name "Radicchio", which has been adopted by all common languages, indicates a very differentiated group of chicories with red or variegated leaves that are traditionally cultivated in North-Eastern Italy.

Radicchio is considered a promising crop, as it is colorful; indeed, Radicchio is added to the color variety of fresh-cut mixed salads, for which there is an increasing demand by consumers [26]. There is no documented history about the origin of colored leaf chicory in Italy. All red types of Radicchio currently cultivated appear to derive from red-leaved individuals first introduced in the 15th century. According to Bianchedi [25], the cultivation of red chicory dates to the first half of the 16th century. It is clear that the original type should be identified with "Rosso di Treviso Tardivo (Late Red of Treviso)" which was long the only cultivated Radicchio in the Venetian territories. After spreading to nearby territories, the original type underwent accentuated selection according to very different criteria suggested by personal preference, but at least partially due to or dependent on the various environmental situations. Thus, after years of repeated selection and intercrossing by farmers between individuals inside the population tending to produce a heavy head with imbricated leaves, a new type called "Rosso di Treviso Precoce (Early Red of Treviso)" became popular. It is worth noting that the adjectives "late" and "early" do not refer to the cycle but instead indicate two different varietal types. As a matter of fact, there are early and late varieties (this characteristic is usually indicated in days from transplanting) of both types. Similarly in the area of Verona, a small winter hardy type forming a rosette of deep-red colored leaves with a nice egg shape was initially selected from "Rosso di Treviso"; from this, the most recent populations of "Rosso di Verona" (Red of Verona) were obtained around 1960. Currently, both winter varieties and also very early ones belong to this type, which are also suitable for September harvesting and spring production. Later, possibly due to spontaneous or controlled crosses between red-leaved individuals and plants of *C. endivia*, types with red spotted or variegated leaves were originated; these are currently known as "Variegato di Castelfranco" (Variegated of Castelfranco, a small medieval country in the province of Treviso). In the area of Chioggia, a traditional horticultural area established on the sandy soils extending southward from this small sea-side town just south

of Venice, a variegated type able to form rather conical, firm, and tightly closed heads was generated most likely from the Variegated of Castelfranco; in the field, the new variety was originally selected circa 1930. From this variety, a large-leaved red type with an accentuated and white midrib and characteristic ball-shaped heads, known as “Rosso di Chioggia” (Red of Chioggia), was initially selected approximately twenty years later, and an almost completely light-yellowish type of very limited cultivation named “Bianco di Chioggia” (White of Chioggia) was obtained toward the end of the last century.

As a result, at least five main cultivated types are grown, named according to their province or town of origin, and may at present be distinguished within this cultivar group (**Figure 1**).

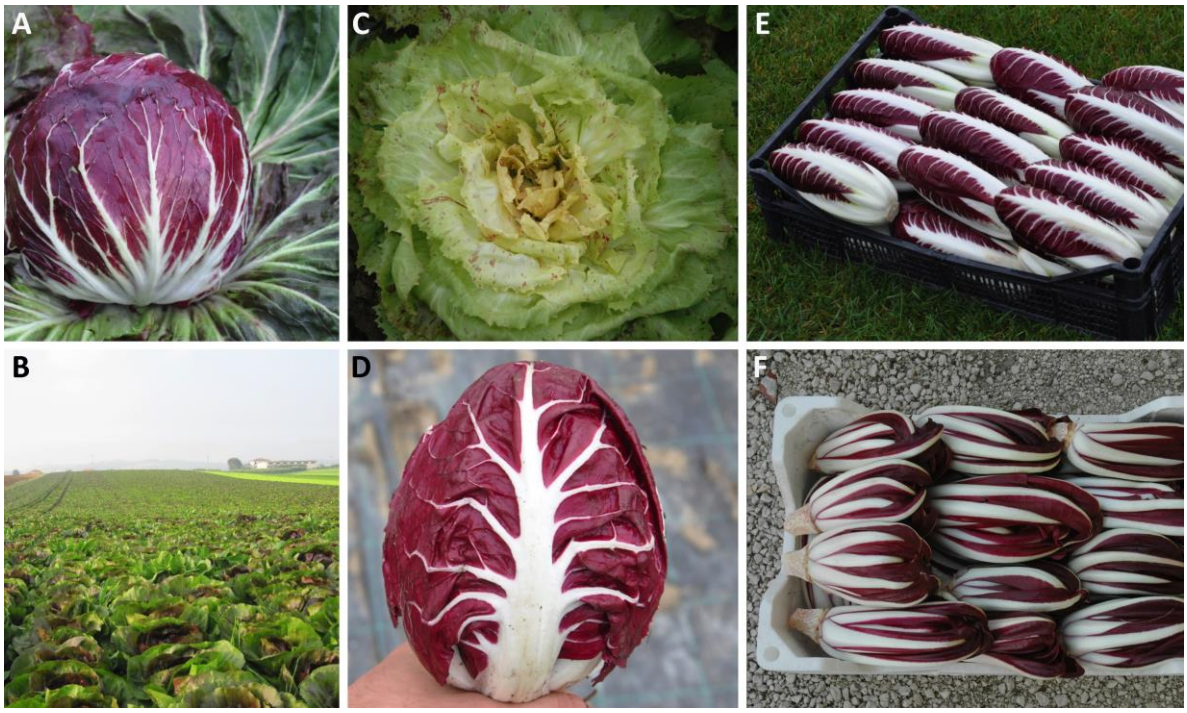


Figure 1. An overview on the main cultivated biotypes of Italian Radicchio: “Rosso di Chioggia” (Red of Chioggia, panels A-B); “Variegato di Castelfranco” (Variegated of Castelfranco, C); “Rosso di Verona” (Red of Verona, D); “Rosso di Treviso Precoce” (Early Red of Treviso, E) and Rosso di Treviso Tardivo (Late Red of Treviso, F).

The variety “Red of Chioggia” is by far the most widely grown among the various types of Radicchio and variety that presents the highest within-type differentiation among cultivars able to guarantee production almost year-round. Indeed, Rosso di Chioggia has exhibited great adaptability to very different environmental situations worldwide, becoming the most grown type of Radicchio outside Italy and thus the most known at the international level [2]. Independent of the sowing time, it grows in the open field. The great majority of farmers in the typical area of production still use seed from their own populations, which they maintain through yearly conservative selection and on-farm seed production. This seed is often sold through private transactions, outside the official seed market, both inside and outside the typical area. Commercial seed is mainly utilized outside the typical area of cultivation. The majority of commercial varieties are open-pollinated populations derived through selection from the original genetic pool. In recent years, both synthetic and hybrid varieties have been put on the market and have been favorably adopted largely for out-of-season production.

The variety “Early Red of Treviso” is characterized as having upright long leaves with a large and thick midrib sustaining a rather expanded deep-red-colored lamina. During the vegetative period, the newly developed leaves do not expand in an open rosette as the plant grows but tighten to form closed and firm heads. In the Veneto region, it is sown or transplanted in the field from July to mid August and harvested in September through December-January. At harvest, the outer green leaves and the major part of the taproot are removed to leave the inside red heart ready for market. Although it is one of the most recent selections, it is becoming increasingly known outside its initially limited area of production. Due to a cultivation technique very similar to the one used for “Red of Chioggia”, this type of Radicchio is beginning to show the same trend of expansion. As a consequence, the seed industry has been looking with increasing attention at the “Early Red of Treviso”. In addition to some open-pollinated commercial populations, some hybrid varieties have been put on the market. Because out-of-season cultivation has started to be adopted, an increasing demand for genetically improved material is foreseeable. At present, most farmers utilize their own populations derived from the original genetic pool and maintained through yearly mass selection. In recent years, synthetic varieties released by

private firms and pre-commercial hybrid varieties have been released on the seed market, as occurred ten years prior for Rosso di Chioggia.

Among the cultivated biotypes, the “Late Red of Treviso” is the most ancient type of Radicchio grown in Italy and can be considered the legitimate ancestor of all others. It is a typical winter product, as it is sown or transplanted in the field from July to mid-August and may be harvested in October through February. The plant develops long, deep-green, essentially upright leaves that form a loose rosette; the midrib and lamina produce an always more accentuated reddish color as temperature decreases. At harvest, the entire plants are dug up, stored with all their leaves and roots intact, and maintained at low temperature, approximately 0°C, as long as possible. According to the market’s request, plants are forced by placing them under a black cover, with their roots in running water, at 10-12°C. After 10-18 days, the forcing period is concluded according to the air temperature; the plants are cleaned, whereby the outer leaves and a great portion of the tap root is eliminated, leaving a bunch of bright-red colored leaves with a white large midrib and a rather reduced lamina. This crop has at least some features in common with Witloof chicory, with which it shares leaf shape, growth habit, a large taproot, and a forcing process to obtain the commercial product. It is grown in a very restricted area and, together with “Variegated of Castelfranco”, “Early Red of Treviso Precoce”, “Red of Chioggia” and “Red of Verona”, is one of the five Radicchio types recognized since the late ‘90s with the PGI (Protected Geographical Indication) mark. Its peculiar aspect and rather superior culinary quality make this Italian Radicchio the most appreciated type. Its market price, particularly at Christmas, can reach as high as twice or three times the price of any other Radicchio. In spite of this, no named commercial variety is on the seed market except for selected open-pollinated populations. In fact, its cultivation is very limited, and the entire production procedure is rather complicated and much less standardized than that adopted for Witloof chicory. Therefore, unless it reaches a comparable degree of popularity, it seems unlikely that the seed industry would invest in this very unique crop. Almost all of the actual production thus relies on farmer populations, the history of which may extend for generations and which are maintained via yearly mass selection.

Concerning the “Red of Verona”, the first populations of this type of Radicchio were obtained approximately sixty years ago. With respect to previous populations, the actual

populations comprise plants with much larger heads that, although on average smaller, may resemble those of “Red of Chioggia”. In comparison, the “Red of Verona” heads are more egg-shaped, formed by less expansive leaves with a brighter red lamina and a large and thick midrib from which less evident and intersecting veins extend. The cultivation period of this type is reduced compared to that of “Red of Chioggia” and “Early Red of Treviso”. It is a typical winter crop, the popularity and cultivation area of which are increasing both in Italy, where it is extending to more southern regions, and outside the country. The reason for this expansion is much the same as for early “Red of Treviso”. Indeed, its cultivation can be standardized quite easily and there is no need for forcing. Moreover, due to its attractiveness, the product is well accepted by the market, and the consumer recognizes a better culinary quality in comparison to other Radicchio types. In spite of this increasing popularity, the seed market is rather poor, and available commercial varieties are selected open-pollinated populations. The first hybrid variety appears to be forthcoming. The most frequently used seed is thus from farmer populations selected during the last decades and maintained through mass selection. It is worth noting that in developing these populations, a procedure suggesting crosses of the initial small-leaved “Red of Verona” with the larger headed type “Red of Chioggia” may have been adopted. Regardless, growers appreciate the type that produces large heads called “Cologna” because it allows higher yields, whereas the marker prefers the original small “Verona”.

Together with “Late Red of Treviso”, the “Variegated of Castelfranco” is the second most traditional type of Radicchio grown in north-eastern Italy. Its morphological traits make it easily distinguishable from any other type. Directly sown or transplanted between July and the end of August, the plants develop a large rosette of more or less indented brilliant green leaves with a very extended red-spotted or variegated lamina sustained by a not-too-evident white midrib. At harvest, the external green leaves are removed, and the internal leaves open to form a bunch of creamy-yellowish red-spotted leaves that look very much like a flower. Indeed, this Radicchio is also known as “Rose of Castelfranco” and is one of the most appreciated components of fresh salads during the cold season. The selected populations grown at present are all self blanching; thus, cultivation of “Variegated of Castelfranco” is comparable to that described for other types of Radicchio, with the exception of “Late Red of Treviso”, for which a final forcing process has long been used.

The availability of commercial seed is much the same as for the majority of other Radicchios. Selected populations or synthetic varieties are available on the market, but the great proportion of the crop is planted using seed of farmer populations long selected and maintained through mass selection by each farmer.

Genetics and Breeding of Radicchio

C. intybus (cultivated types) is a biennial species, while wild chicories are perennial. Although there are differences according to the cultivar group, early sowing or transplanting in spring under long days results in almost generalized flowering. Conversely, if sowing or transplanting are delayed until July and August, the plant forms a rather loosen rosette or a fairly compact head that remains in the field until the following spring; at this time, between May and June, the central bud develops into a stem bearing inflorescences with clusters of blue flowers (rarely petals white or mauve).

Many clusters of 2-4, rarely 8, sessile "flowers" are developed on the flowering stalk in axillary position or single inflorescences at the end of peduncles 4-7 cm long, rarely up to 13. The inflorescence (capitulum) is typical of the entire family and consists of a cluster of 15-25 single hermaphrodite flowers on a receptacle that is protected by an involucre. Each single flower has a gamopetalous and ligulate corolla and bears five filamentous stamens fused by their anthers to form a column surrounding a pistil with a bifid stigma.

At flowering, the style elongates, and the stigma is pushed up through the small channel composed of the anthers; the two halves of the stigma separate and assume a rather pronounced spiral form that may bring the inner receptive surface, completely free from pollen, in contact with the outer surface of the pistil, which, extruding from the staminal column, has remained densely covered with pollen grains. Thus, independently of the intervention of external agents, self-pollination is possible in addition to open-pollination. In general, *C. intibus* is characterized by a strong sporophytic incompatibility system that usually inhibits self-fertilization and hence favorites cross-fertilization [2-10-26-27-28-29-30]. Flowering occurs in the early part of the morning, mainly between 7:00 and 10:00 a.m., according to the air temperature and humidity. In the afternoon, after 01:00 p.m. the flowers wilt. The flowers can remain turgid and the pollen viable until the first hours of the afternoon only under exceptionally cold and humid conditions or during cloudy days.

Productive as well as qualitative traits are the main objectives in Radicchio breeding programs. The general and common goals of breeding new varieties mainly concern the following: i) single plant size, weight and yield; ii) resistance to biotic (fungal diseases and insects) and abiotic stresses; iii) adaptation to specific climatic or agronomic environments;

iv) uniformity of crop maturity, size and growth; v) good market acceptance as far as extrinsic (color, shape, uniformity) and intrinsic (taste and texture) traits.

There is 90 years of breeding history for *C. intybus*, dating to when the first varieties were bred and sold on the seed market. With regard to Radicchio, the main traits evaluated during selection programs are related to morpho-phenological, agronomic, and organoleptic characteristics. Important features are the time of cultivation, class of earliness, thickness and length of the main root, leaf shape and color, adaptation to local environments, disease resistance, and taste and bitterness of the edible parts.

Regardless, the performance of a cultivar is strictly dependent upon its genetic value, which in turn may be tightly linked to its genetic structure and thus to the strategy adopted for its constitution. Traditionally, varieties of Radicchio were developed by mass selection to obtain uniform populations characterized by valuable production and acceptable commercial head size and shape. Newly released varieties are mainly synthetics produced by intercrossing of a number of phenotypically superior plants selected on the basis of morpho-phenological and commercial traits. More rarely, plants are also evaluated genotypically by means of progeny tests. Synthetics have a large genetic base and are represented by a heterogeneous mixture of highly heterozygous genotypes sharing a common gene pool.

In recent years, methods for breeding F_1 hybrid varieties have been developed by private breeders and seed companies. However, details are not available in the current literature, and it may be presumed that each company has developed its own protocol, mainly in accordance with the genetic material it has at its disposal and the possibility of applying more or less efficient control at the F_1 hybrid seed production phase. In fact, the strong self-incompatibility system, which makes it difficult to obtain highly homozygous parents, has always been considered to be a great barrier for the production of inbred lines and parental lines for the development of hybrid varieties. Increased interest in the production of F_1 hybrid varieties has resulted from the discovery of spontaneous male-sterile mutants [31-32]. Male-sterility, or the inability of a plant to produce functional pollen, is important for the commercial production of hybrid seed via the crossing of parental inbred lines appropriately selected through progeny tests; such new commercial varieties can be considered to be true F1 hybrids.

Chicory breeding materials of the Radicchio group are usually represented by local populations known to possess high variation and adaptation to the natural and anthropological environment where they originated and are still cultivated [33]. These populations are maintained by farmers through phenotypic selection according to their own criteria. Controlled hybridization among different types is occasionally exploited to obtain recombinant genotypes showing superior agronomic and commercial traits or to avoid inbreeding depression caused by yearly intercrossing among selected individuals of the same population to improve uniformity. The ongoing breeding programs by local breeders and regional seed institutions aim at the following: i) to isolate, within the best local selections, individuals amenable for use as parents for developing synthetic varieties, and although not easily feasible; ii) to select inbred lines suitable for the production of commercial F_1 hybrids. These breeding procedures could be greatly assisted by the use of molecular markers that allow the discarding of molecular off-types to better exploit the parental genetic polymorphisms for synthetics and to identify the most genetically distant inbreds as parental lines for hybrids.

A scheme widely used for the constitution of synthetics in all Radicchio types includes the field selection of approximately 50-100 mother plants conforming to a prefixed breeder's ideotype. These plants should share similar morpho-phenological traits, and productive and qualitative properties. Local varieties and landraces, and especially populations from repeated cycles of mass selection, represent excellent breeding material for the selection of the mother plants. In addition to morpho-phenological, agronomic and commercial evaluation, characterization based on molecular markers could be undertaken to choose and retain plants of the original population that associate a high genetic uniformity with a superior phenotypic value. The roots of the selected mother plants should be preserved during the winter until the next spring, when they will be transplanted in the field under an isolation cage. The seed from plant intercrossing is the basic stock from which commercial seed lots can be obtained and sold to farmers. It is generally accepted that seed multiplication should no longer be performed in the cultivation areas of Radicchio. Traditional dedicated areas for seed production are the Apennine valleys of central Italy, which allow the production of high-quality seed (*i.e.*, germinability and purity) due to the favorable climate.

The application of molecular markers to Radicchio breeding, as has been achieved with other crops for which they have been extensively and routinely utilized, may be of great help in overcoming theoretical as well as practical problems. Solutions to these problems can improve the efficiency of possible breeding methods and schemes.

Phylogeny, population distinctiveness and their interrelationships as well as the genetic traceability of the commercial product are only a few examples of the most frequent uses. For instance, the possibility of discriminating the five different types of Radicchio grown in Veneto (Italy) was established on this basis [33]. More refined applications concern the construction of genetic linkage maps, the location of specific loci for genetic factors affecting important qualitative traits and the possibility of quantifying their contribution to the total explained variation of quantitative characters. The mapping of useful genes should facilitate the identification of superior genotypes and minimize the manifestation of linkage drag effects in backcrossing strategies. Most importantly in Radicchio, the implementation of marker-assisted breeding programs based on mapped molecular markers could be very useful for the selection of parental genotypes of synthetic and F_1 hybrid varieties.

Molecular markers could also be useful to test other important varietal attributes in synthetics, such as mother plant relationships, heterozygosity evaluation and prediction, population uniformity and distinctiveness, and to assess the homozygosity and stability of inbred lines and diversity among inbred lines, in order to maximize heterosis in hybrid varieties. A cost- and time-efficient multilocus genotyping method based on a panel of mapped SSR markers is now available for Radicchio [34]. Its application is having utility for assessing the degree of homozygosity of parental inbred lines, as a measure of their genetic stability, and also for predicting the degree of heterozygosity of their F1 hybrids, as an estimate of the specific combining ability on the basis of the genetic diversity between maternal and paternal inbred lines.

The Reproductive System of Chicory

In *C. intybus*, the mating system is usually characterized by mechanisms that promote cross-fertilization and reduce or impede self-fertilization. In most cases, this is directly related to the manifestation of inbreeding depression and heterotic vigor as result of selfing and outcrossing, respectively [30-35-36].

C. intybus is a self-incompatible species. Both auto-pollen and allo-pollen grains deriving, respectively, from the same genotype or different but related genotypes are recognized by cells of the pistil and rejected after pollination, either immediately on the stigma surface or later during pollen tube growth in the transmitting tissue of the style. Genes located at the S-locus encode the male (pollen) and female (pistil) recognition determinants of self-incompatibility (SI).

Over 60% of Asteraceae are estimated to be self-incompatible with sporophytic genetic determination [37-38]. The existence of an SSI system in chicory was demonstrated in two different studies analyzing the progeny of two crosses between inbred lines of Witloof chicory [26] and the progeny of the crosses between a wild-type chicory plant with a cultivated Radicchio plant of "Red of Chioggia" [27]. In both studies, di-allelic crosses among the F_1 plants and F_1 backcrosses with both parents were applied to analyze the progeny for incompatibility reactions. The mean number of achenes per inflorescence was chosen as the criterion for evaluating the compatibility or incompatibility of a cross. The results indicated that a one-locus sporophytic self-incompatibility system is present in chicory, and the existence of dominance and co-dominance relationships between *S*-alleles in the pollen and styles was noted. In *Brassica*, it has been shown that dominant interactions between *S*-haplotypes can act independently in pollen and stigma, thereby resulting in extremely complex patterns of incompatibility among individuals [39]. This also appears to be the case in *Cichorium* because only a difference in pollen and stigma allelic activity can explain the seed set observed in certain cross-combinations [27]. Further studies [2-10-29] confirmed the presence of a sporophytic self-incompatibility system in the chicory.

Indeed, different levels of self- and cross-incompatibility are found when wild and cultivated genotypes are crossed [27-40]. It has been suggested that the observed self-compatibility could be due to the presence of heterozygous genotypes bearing *S*-alleles with

stigma and pollen dominance relationships that are not linear and the presence of homozygous genotypes with weak *S*-alleles [41]. Moreover, recessive *S*-haplotypes can reach high frequencies in chicory populations because their effects are masked by dominant *S*-haplotypes. It should be underlined that in spite of the SSI mechanism, self-pollinated Radicchio plants very often yield one or a few fertile seeds per flowerhead, which means that considering the large number of flowerheads produced by a single plant during its flowering period (8-10 weeks), the seed set can be significant and this make possible the production of inbred lines.

Rejection of incompatible pollen on the chicory stigma is very rapid in the case of self-pollination because auto-pollen grains do not adhere to papilla cells. In incompatible crosses, papilla cells frequently permit the development of the pollen tube, which bursts and does not reach the transmitting tissue of the style. The rejection process in *Cichorium* differs from that of *Brassica* because the incompatibility response manifests within a few minutes due to the inhibition of pollen hydration or pollen germination or pollen invasion of the stigma epidermis [42].

Few data are available on the molecular mechanisms operating in chicory or in other species belonging to Asteraceae. In Radicchio, molecular investigations have sought to identify orthologs of *SRK* and *SLG*, the female components of the *Brassica S*-locus. The *SRK*-like genes in chicory were found to not be exclusively expressed in stigmas, thus indicating that they are unlikely candidates for stigma *S*-genes [43]. Analyses of stigma proteins in *Senecio squalidus* have revealed a family of polymorphic basic proteins associated with specific *S* genotypes, though these proteins bear no resemblance to the *S*-locus product of *Brassica* spp. [39-44]. It will be useful to ascertain whether Asteraceae species possess their own system of self-incompatibility at the molecular level, as further information is needed for manipulating the SI system in chicory breeding.

High-density genetic maps including the loci responsible for sporophytic self-incompatibility (*S*-locus) and involved in nuclear male-sterility (*NMS-1* locus) have been published [30]. This study was performed using a mapping population of 389 F_1 individuals obtained by crossing two plants, one male-sterile and one male-fertile, both heterozygous at the *S*-locus, according to a method described by Eenink [26] and Varotto *et al*. [27]. Phenotyping of progeny plants for the male-sterility or -fertility trait allowed mapping of

the *NMS1* locus to linkage group 5, and controlled diallelic and factorial crosses were utilized to identify self-compatible or -incompatible phenotypes for mapping the *S*-locus to a 0.8 cM region on linkage group 2. A bulked segregant analysis was performed using the AFLP technology and a total of 2,350 out of 31,000 markers were found to be polymorphic between parents and segregating in the progenies. Thirteen and six AFLP markers were found genetically linked to the *NMS1* locus and the *S*-locus, respectively. Eight of these markers were converted into locus-specific SCAR diagnostic assays, among which five exhibited co-dominant polymorphisms [30].

It is worth mentioning that the chromosomal blocks containing the genomic loci for male-sterility and self-incompatibility were both encompassing a region of less than 1 cM in length. In addition, Gonthier *et al.* [30] mapped genes encoding proteins similar to *S*-receptor kinase, the female determinant of sporophytic SI in Brassicaceae, and markers related to the putative *S*-locus of sunflower, family of Compositae, but none of these genes or markers mapped close to the chicory *S*-locus. Concerning male-sterility, neither candidate genes nor diagnostic markers are available for *Cichorium* spp.

Male-sterility may have an important function for breeding F1 hybrids of leaf chicory, particularly because the use of SI for the female line is inadequate for reliable F1 hybrid seed production on a commercial scale. Male-sterility is another efficient mechanism for outcrossing in angiosperms and is represented by mutants that cannot produce pollen or fertile pollen grains, functional anthers or gametes, though female-fertility is normally preserved [45-46]. Two types of male-sterility can be observed in plants, which can be divided into nuclear (NMS) and cytoplasmic male-sterility (CMS) on the bases of inheritance patterns. Nuclear male-sterility is based solely on recessive mutations that affect different functions of nuclear genes, whereas the cytoplasmic type is maternally inherited and mainly due to mutations in mitochondrial genes [47]. Moreover, male-fertility can be restored in CMS genotypes by nuclear-encoded fertility restorer (*Rf*) genes. In fact, CMS has been used in several species to produce female parental lines and has also been used for the production of hybrid seeds. In the absence of this system, male floral organs must be removed mechanically.

CMS is preferred by breeders over NMS owing to its modes of transmission to progeny and restoration of fertility [47-48]. Regardless, the presence of a naturally occurring CMS

system has not been reported in chicory [2], whereas two spontaneous NMS mutations have been discovered and exploited by breeders: one in root chicory [49-51] and the other in leaf chicory belonging to the biotype "Red of Chioggia" [31-32].

With regard to the patents by Barcaccia and Tiozzo [31-32], the invertors described the mutants at the cytological and genetic levels, affirming that the mutation itself affects a single nuclear gene, providing a recessive trait that causes male-sterility when homozygous. Moreover, the inventors documented that the mutation caused the microspores of each tetrad to arrest development at the uninucleate stage, degenerating before their release from the tetrads. At full flowering in the male-sterile genotypes, all the microspores of dehiscent anthers were found to be shapeless, shrunken and much smaller than those of wild-type. The inventors demonstrated that pollen grains are never produced in mature anthers, supporting full expressivity of the trait, with mutants being 100% male-sterile. The gene *ms1* responsible for male-sterility was mapped at 5.8 cM upstream and 12.1 cM downstream two SSR markers [31-32]. This information enabled to develop a first diagnostic assay useful to discriminate male-sterile and male-fertile plants of Radicchio, with a low genotyping error. The gene *ms1* has been recently mapped on linkage group 4 [50] into a chromosomal region spanning 8.5 cM (**Figure 2**).

As reported by Cadalen *et al*. [50], a MADS-box gene is present a few centiMorgans from the SSR marker loci mapped by Barcaccia and Tiozzo [31-32] on linkage group 4. Functional analyses by molecular genetic studies in model eudicots, such as *A. thaliana*, have shown that transcription factors encoded by MADS-box class L2R2 genes are essential for regulating various aspects of flower development [52-53]. Further investigations on this gene were performed in chicory by Collani [54] on the basis of the findings in rice [55]. Both gene expression and segregation data have shown that this MADS-box candidate cannot be the genetic factor responsible for male-sterility in Radicchio.

Approaches to genetic engineering of male-sterility have been applied to Magdeburg, Witloof and Chioggia genotypes [56]. Nevertheless, this technique is currently not in use, as the public opinion disapproves of the application of GMOs in food production. Moreover, the absence of joint EU legislation makes the development of such projects not profitable for seed companies.

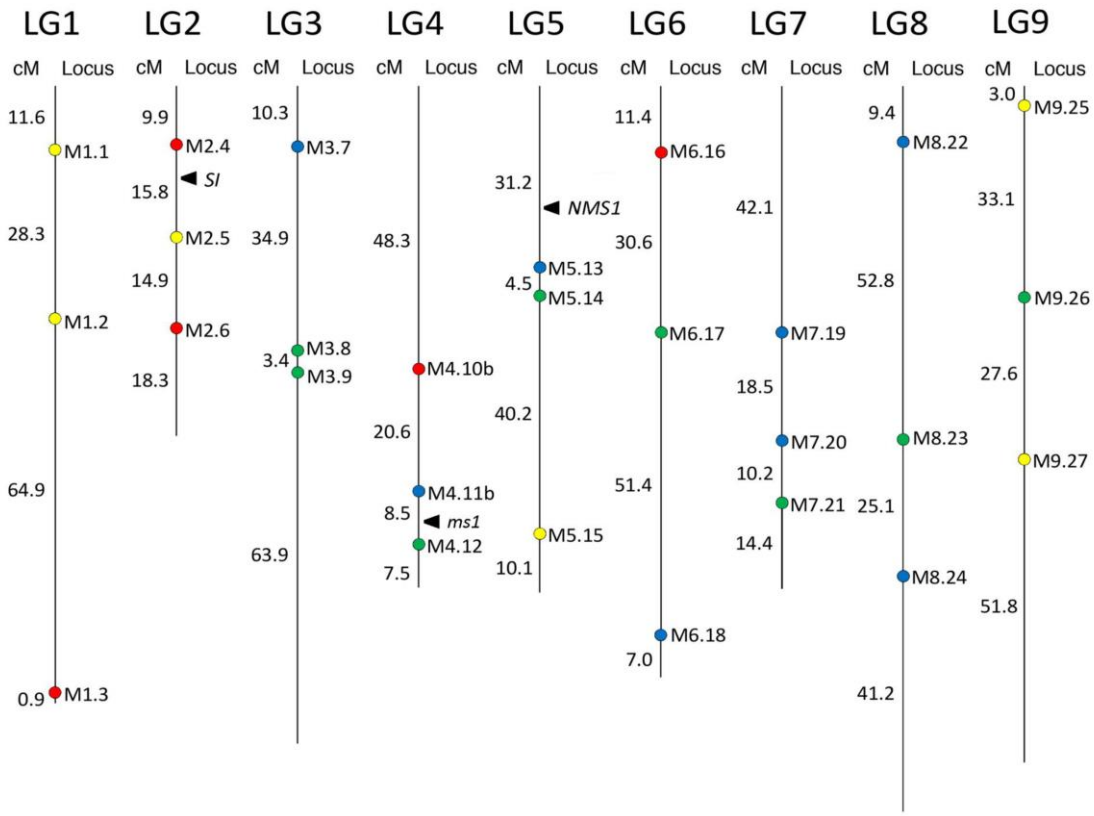


Figure 2. Schematic representation of the 9 basic linkage groups of chicory (*Cichorium intybus* L.) with indication of the 27 mapped marker loci chosen by Ghedina *et al*. (2015) to implement a method for genotyping elite breeding stocks. Mutant genes responsible for genetic male-sterility were mapped on linkage group 5 (*NMS1*-locus) and 4 (*ms1*-locus) whereas the *S*-locus controlling self-incompatibility was mapped on linkage group 2 (*SI*). For details see Cadalen *et al*. [50], Gonthier *et al*. [30], and Barcaccia and Tiozzo [31-32].

Another method pursued by different groups [57-58] is protoplast symmetric fusion between a breeding line of chicory and a CMS line of sunflower, which allowed regeneration of interspecific hybrid plants. To transfer the male-sterile cytoplasm from an industrial Chicory nuclear environment into a Witloof background, three different CMS cybrids originating from three different fusion events were characterized at the molecular level and then backcrossed to Witloof [59].

Furthermore, protoplast asymmetric fusion was used to produce male-sterile somatic hybrids between a Radicchio accession and a PET-1 sunflower CMS line [60].

Other works in sunflower have focused on male-sterility and have led to the identification of several NMS genes as well as CMS and associated restored genes, some of which have been finely mapped [61-62-63]. On the basis of these findings, Habarugira *et al*. [64]

reported the effects of the nuclear genome on anther development using a chicory CMS mutant. These researchers studied histo-morphological changes during anther development in fertile chicory plants from flower initiation to anthesis. Moreover, using normal development as a reference in fertile plants, alterations in the male-sterile mutant named "524" (*i.e.* a male-sterile cybrid obtained by protoplast fusion) and in different phenotypes obtained from crosses between this male-sterile plant and two different pollen donors were investigated [64].

The genetic relationships among the species and cultivar groups of *C. intybus* were established by Kiers *et al.* [10] using molecular markers and multivariate statistics. At the species level, the results correspond with previously obtained phylogenetic relationships, with *C. bottae* as the most divergent species and *C. intybus* and *C. spinosum* as well as *C. endivia*, *C. pumilum* and *C. calvum* clustered into distinct subgroups. Based on the congruence between phylogenetic and genetic analyses, unique markers were expected for all species, except for *C. bottae*. The analysis of *C. intybus* materials resembled the species analysis in terms of sorting cultivars according to cultivar groups. In contrast to *C. intybus*, the cultivar series of *C. endivia* did not form distinct groups, suggesting that crosses have been made among the various cultivar groups. These findings were recently confirmed by Raulier *et al*. [18].

Chicory (*C. intybus*, 2*n*=2*x*=18) is a strictly allogamous species being strongly hampered by an efficient incompatibility system, which prevents inbreeding and promotes outbreeding [26-27-28-30].

Regarding their genetic structure, the original populations of C. *intybus* could be considered as natural because independently of their historic background, the production of both Witloof and Radicchio has long relied on populations maintained by farmers for their own use. In these cases, very little selection, if any, might have been applied according to personal criteria. All these populations obtained by mass selection and maintained through the intercrossing of selected phenotypically superior mother parents can be considered as highly heterozygous and genetically heterogeneous, with behavior and level of adaptation to different environments and/or cultural conditions dependent on the frequency of favorable genes or gene combinations [2]. As interest in edible products grew, the selection criteria of farmers became increasingly attentive to the consumer's and market's demands, and most of them elaborated their own ideotypes. This situation brought about a great deal of genetic and morphological differentiation that was entirely preserved until organized breeding programs were implemented, first by public institutions and in more recent times by private firms. As occurs for most open-pollinated species, detectable heterosis (*i.e.*, hybrid vigor) effects are present in *C. intybus*. Hybridization between selected genotypes provides uniform and heterotic progeny due to increased heterozygosity. Thus, the

development of F_1 hybrid varieties is feasible. Except for the "Red of Chioggia" and "Early Red of Treviso" biotypes, for which both synthetic and hybrid varieties have been released onto the market by private firms, the vast majority of the Radicchio crop is still based on farmer's populations. These populations are selected and maintained annually, and the seed is usually reutilized on the farm but it may also be sold through private and non-officially registered transactions. Usually single populations are very well distinguishable among types, being very often also recognizable within a given type on the basis of morphological and physiological characters and agronomic performance. At the same time, they present an acceptable phenotypic uniformity among individuals. With respect to genetic variation estimated by applying molecular markers, a common observation true for both Radicchio and Witloof chicories is that the vast majority of the genetic variation can be explained at the within-population level, whereas only a minor portion is attributable to among-population differences [15-19-33].

In 2015, two different groups published two independent works on the genetic characterization of leaf chicory breeding stocks based on data collected by exploring mapped SSR markers.

The data presented by Raulier *et al.* [18] agree with previous studies. Utilizing 15 SSR markers, the researchers analyzed the genetic diversity of the current industrial chicory germplasm and obtained data concerning the relationships between and within *C. intybus* L. and *C. endivia* L. A total of 19 cultivated *C. endivia* lines, along with 27 wild and 155 cultivated *C. intybus* populations were analyzed, corresponding to 42 Witloof, 83 root chicory and 30 leaf chicory accessions (the latter included Radicchio, Sugarloaf and Catalogne materials). Moreover, 1,297 samples derived from 15 modern root chicory cultivars were also processed. The authors reported that the *C. endivia* and *C. intybus* accessions could be clearly separated from each other. However, they also confirmed that seven wild *C. intybus* individuals were genetically closer to *C. endivia* than to *C. intybus*, showing the complex genetic interrelationships between these *Cichorium* species. Furthermore, based on this analysis, the authors confirmed the existence of three cultivar groups of *C. intybus* (*i.e.*, Witloof, root and leaf chicory) and the three subgroups of leaf chicory (*i.e.*, Radicchio, Sugarloaf and Catalogne). For industrial root chicory cultivars, evident differentiation was observed among Belgian, Polish and Austrian varieties, whereas

no differentiation was detected among the French modern varieties. Overall, the researchers reported that most (*i.e.*, 87%) of the genetic variation occurs within single cultivars, with 7% among cultivars of the same biotype and 6% among cultivars of different biotypes.

In the same year, Ghedina *et al*. [34] published a method for genotyping elite breeding stocks of leaf chicory by assaying 27 mapped SSR markers, three per linkage group. As the plant accessions analyzed were inbred lines, the data were not comparable with the data of previous studies, in which open-pollinated varieties or wild accessions were taken into account. The authors demonstrated that most of the genetic differentiation occurs among populations, as expected, and that each population deriving from repeated selfing cycles can be considered as being genetically uniform and distinguishable from the others of the core collection of Radicchio breeding materials. Nevertheless, they also observed that inbreeding coefficients of single lines were low or negative, indicating that the observed heterozygosity is higher than expected. An interesting hypothesis is that maintenance of such levels of heterozygosity despite inbreeding reproductive strategies (*i.e.*, selfing, full-sibling and back-crossing) could be a consequence of the reproductive system of *C. intybus*, which is naturally characterized by a high frequency of allogamy as a result of self-incompatibility [34]. In addition, the authors also speculated that a fraction of the observed heterozygosity could be a consequence of phenotypic selection of morphologically superior individuals performed by breeders during inbreeding programs.

Because the local farmer's seed still nowadays represents the starting materials for the development of new commercial varieties, it seems reasonable to state that, if preserved from extinction, these local varieties are an invaluable germplasm source on which leaf chicory breeding programs can rely for a long time. Modern varieties of "Red of Chioggia" and "Early Red of Treviso" are mainly synthetics and F1 hybrids, the latter being bred and adopted with increasing frequency, as predicted by Lucchin *et al*. [2] This is particularly true for those biotypes that can take an advantage of the uniformity of the marketed product, as this is often the key to customer's appreciation.

Although chicory cannot be considered as a model species, some genetic studies using molecular markers have been conducted in order to characterize commercial varieties and experimental materials [20-33-65-66], to evaluate the genetic homogeneity and purity of inbreds and hybrids, respectively [34-67], to investigate phylogenetic relationships between

cultivars and cultivar groups of *C. intybus* and other species, both cultivated and wild, belonging to the same genus [12-15-18-19-68] and to assess the degree of spontaneous gene flow between landraces, cultivars and wild populations of chicory. Moreover, molecular markers were also applied to understand whether co-existence between transgenic varieties and conventional varieties is possible [69].

Molecular marker-based investigations have also been aimed at evaluating the genetic relationships among the five types of Radicchio and establishing a molecular reference system that would allow the precise identification of the different cultivated types. The five major cultivated types of Radicchio were investigated by PCR-derived markers [33]. The experimental material was represented by two outbred populations (one of Variegated of Castelfranco and one of Red of Verona) and by eight partial inbred lines (three of Early Red of Treviso, three of Late Red of Treviso and two of Red of Chioggia). The different types were well distinguished from one another when analyzed in bulks using AFLP markers at the population level. However, they were not well distinguished when analyzed at the individual level using RAPD, inter-SSR and AP-PCR markers. The genetic variation was shown to be much higher within types than between types. This result suggested that in each Radicchio type, populations produced by breeders through controlled intercrossing or repeated selfing have maintained good separation of gene pools over the years.

On the basis of the reproductive system of Radicchio, such a finding may be explained by taking the following into account: i) the SSI system of the species, which -limits both selfing and intercrossing between plants with an identical phenotype at the multi-allelic *S*-locus, thus allowing a certain amount of heterozygosity to be maintained even in inbred populations; ii) the selection criteria of mother plants applied each year by each farmer to maintain his/her own population most likely results in limited contamination between types, thereby preserving the phenotypic identity of each type. At the same time, seed multiplication that has been carried out over the years has produced a clear genetic differentiation between types within the species [33].

The establishment of a molecular reference system in Radicchio appears to be feasible and suitable both for the precise identification of single biotypes and for evaluating of the extent of the natural hybridization that can occur between different biotypes. Diagnostic molecular markers, along with morphological and phenological descriptors, will be useful for the

certification of typical local products of Radicchio and for the recognition of a protected geographical indication (PGI) mark [2].

In the last two decades, molecular markers have been exploited for the construction of linkage maps and for the identification of Mendelian genes and QTLs. In particular, AFLP and RAPD markers were used to construct the two oldest genetic linkage maps of *C. intybus* var. *latifolium* reported in the literature [21-70]. More recently, Cadalen *et al*. [50] used SSR and STS markers to develop a consensus genetic map for this species containing 472 marker loci and covering 878 cM. This genetic map was integrated by Gonthier *et al.* [30] for marker loci involved in nuclear male-sterility and sporophytic self-incompatibility using five SCAR markers showing co-dominant inheritance. An independent work conducted by Muys *et al*. [71] enabled to implement a genetic map of industrial chicory (*C. intybus* var. *sativum*) that includes 237 marker loci attributable to AFLP, SSR and SNP markers, and covering 1,208 cM.

Recently, Delporte *et al*. [72] reported data on the selection and validation of reference genes for elucidating the phenylpropanoid pathway and for investigating gene expression patterns and models in chicory. This species accumulates four major polyphenols: chlorogenic, isochlorogenic, caftaric and chicoric acids [73]. These acids, caffeic esters, are well known for their antioxidant and antidiabetic properties [74-75-76]. Although various biochemical pathways are involved in the synthesis of these secondary metabolites, each of them arises from the phenylpropanoid pathway. Eighteen candidate genes were chosen according to known *A. thaliana* stable genes and reference genes commonly used in qRT-PCR studies. Orthologous chicory sequences were retrieved after performing BLASTX searches in a root chicory EST database and specific primers were also designed and tested [72].

An all-encompassing work on the transcriptome assemblies of Compositae crops and wild relatives was conducted by Hodgins *et al.* [68], the main purpose of which was the development of genomic resources for 12 Compositae species, including some of the genus *Cichorium* for which the whole seedlings were used as tissue source (**Table 2**). In this work, both wild and cultivated accessions were represented for the two species of *Cichorium*, *C. endivia* and *C. intybus* [68]. The authors demonstrated a recent domestication for most species of the Compositae family and suggested that all *Cichorium*

taxonomic entities belong to the same biological species because reproductive barriers between crops and wild progenitors are weak or absent. Nevertheless, compared with self-compatible progenitors, self-incompatible progenitors exhibited greater genetic divergence from their corresponding crops. The exception was chicory, which showed the lowest genetic divergence among self-compatible species [68]. The authors also observed evidence of introgression when comparing endive and chicory, a result that was replicated in all four comparisons of the wild and cultivated accessions. Putatively introgressed genes were found and the overrepresented biological processes for these genes were carbon fixation, ATP synthesis-coupled electron transport and mitotic cell cycle spindle assembly checkpoint.

Table 2. Assembly statistics for transcriptomes of four accessions from the *Cichorium* genus (modified from Hodgins *et al.*, [68]). Plant materials include both wild and cultivated taxa of *C. endivia* (*i.e.*, subsp. *pumilum*, wild form, and subsp. *endivia*, cultivated Endive) and *C. intybus* (*i.e.*, subsp. *intybus*, wild form, and var. *foliosum*, Witloof chicory).

Taxon	Reads No.	Contigs No.	Total sequence (Mbp)	Total assembly (Mbp)	Average length (bp)	Max unigene length (bp)
C. endivia (Wild)	18,124,638	52,685	2,157.1	34.0	646	8,355
C. endivia (Endive)	37,847,448	63,647	4,504.4	46.0	723	8,684
C. intybus (Wild)	24,881,864	56,696	2,961.3	38.8	6g84	4,690
C. intybus (Witloof)	22,213,574	58,926	2,643.7	37.4	635	4,583

The first genome draft of leaf chicory was recently unveiled by Galla *et al.* [6]. The plant material used for sequencing of the leaf chicory genome belongs to the Radicchio biotype Red of Chioggia, specifically the inbred line named “SEG111”. This accession was chosen as the most suitable genotype based on five common criteria [6]. Briefly, among the Italian biotypes of leaf chicory, the Red of Chioggia is the most relevant one from a commercial point of view. Plus, the accession used is a male-fertile inbred line that is both phenotypically and genotypically well characterized. It is permanently conserved by *in vitro* culture to allow high availability of clonal materials and it is distinguished by a high degree of homozygosity, more than 80% [34]. Last but not least, it shows an excellent general combining ability, as demonstrated in trial experiments where this inbred line was

used as seed parent in pairwise crosses with genetically distinct pollen donors for the development of F1 hybrids.

Regarding methods, sequencing of the genomic DNA library was performed using Illumina HiSeq and MySeq platforms to combine the high number of reads originating from the former with the longer sequences produced by the latter; the average coverage in a HiSeq run is approximately 21×. The estimated size of the assembled genome draft is 760 Mb. The authors reported to have obtained 58,392,530 and 389,385,400 raw reads using the MySeq and HiSeq platforms, respectively. The two datasets were combined in a unique reference genome draft by assembling 724,009,424 nucleotides into 522,301 contigs. The contig length is ranging from 200 bp to a maximum of 379,698 bp, with an average contig length of 1,386 bp. As much as 68.9% of the recovered sequences are contained within a length spanning 200 nt to 999 nt. The interval length, ranging between 1,000 nt and 2,999 nt, is represented by 19.7% of the assembled contigs, whereas the proportion of contigs with a length longer than or equal to 3,000 nt is 11.5%. The functional annotation of more than 18,000 unigenes was performed according to established computational biology protocols by taking advantage of publically available reference transcriptome data for *C. intybus* [6].

The authors presented the main preliminary findings on the genome organization and gene composition of Radicchio, and the potential of the newly annotated expressed sequences and diagnostic microsatellite markers in breeding programs were critically discussed. Because the main purpose of the sequencing and annotation of the first genome draft of Radicchio was to obtain a powerful tool for breeding programs, and overall statistics on the presence of SSR loci were presented. A total of 66,785 SSR-containing regions were identified. As many as 52,186 and 11,501 sequences were found to contain one or more microsatellites, respectively. The most common SSR elements were those showing a dinucleotide motif (89.0%), followed by trinucleotide (7.1%) and tetranucleotide (3.0%) ones; microsatellites revealing a pentanucleotide and hexanucleotide motif were less than 1.0% of the total.

The functional annotation of the assembled contigs was performed using a BLASTX approach according to which contig sequences were used to query different public protein databases. Data for functional annotations based on the most interesting GO terms in

breeding were presented, with biotic and abiotic stresses, hormonal responses and flower and seed development taken into account. In addition, for each selected metabolic pathway, many microsatellite sequences putatively linked to important traits were presented according to their potential effect on plant characteristics. Based on the Kyoto Encyclopedia of Genes and Genomes database (KEGG, [77]), a total of 22,273 contigs enabled the mapping of 795 enzymes to 157 metabolic pathways that are interesting under the breeding point of view [6].

In conclusion, the availability of the first genome sequence for this plant species can provide a powerful tool to be exploited for the identification of markers associated with or genes responsible for relevant agronomic traits and that influence crop productivity and product quality. As an example, the data and techniques from this research project will be capitalized on in subsequent years for the planning and development of basic studies as well as applied research on male-sterility and self-incompatibility in this species.

Seed Production in Chicory

Seed production represents the conclusive phase of breeding programs and, as with any other cultivated species, is of paramount importance for the success of the crop. In chicory, the seed is the only plant material used for the commercial propagation of varieties and planting, and it often determines the quantity of yield and the qualitative commercial standard of the crop. Despite this, little research has been conducted on chicory seed production. The fruit (seed) is an achene, obovoid to cylindrical, weakly ribbed, and light brown to completely brown when ripe. Almost all of the chicory commercial seed in Europe is produced in The Netherlands, Northern France and Italy, usually in areas with climatic conditions that are sufficiently fair to permit abundant differentiation of flowers, a long flowering period, a good presence of pollinating insects, available water during the filling period, and dry conditions during the last part of the reproductive cycle, when the seeds need to mature and dry to allow a high germination rate (*i.e.*, > 90%). In Italy, for instance, important seed production activity, which is also performed by Italian and European seed companies, is present in the eastern portion of the Po Valley (Veneto region), south of the Po River (Romagna region) and in the Marche region.

Crossing is the rule in *C. intybus*. Different varieties, landraces or, more generally, populations have to be considered as interfertile, *i.e.*, cross-pollination due to visitor insects and cross-fertilization are possible. To avoid this situation and any possible contamination, spatial isolation needs to be strictly respected in different fields (1,500 to 2,000 m) of the seed production areas. Seed producers must also take into account the presence of wild plants, which need to be removed from fallow fields and roadsides. The cultural practices applied to chicory seed crops during the vegetative phase are much the same as those used in growing these plants for the vegetable market, though a lower plant density (approximately 5 plants per square meter, rather than 7-9) is used. During the vegetative period, fields are repeatedly inspected to remove off-type or diseased plants.

As a biennial plant, Radicchio and, in general, all types of leaf chicories require vernalization to differentiate and produce a seed stalk. Thus, a seed crop needs to be sown in autumn to produce seed the next spring. If seed in the same year of sowing is desired, seeding operations need to occur in winter (end of January – beginning of February), ideally under protection, to let the small seed or transplanted plants be naturally vernalized

in the field by the cold temperatures at this time of the year. Another option is to seed in autumn in greenhouses using Styrofoam cell trays and to store the seedlings in cold greenhouses until March. Thus, no losses will be experienced due to cold weather in the field during the winter season. Vernalization will also be assured and flowering will hopefully occur at the same time. However, the genetic control of flower induction and differentiation in chicory has to be better undersood. The transition to flowering was demonstred to be mediated by *MFL* gene [78], homologue of *AtFLC*, the main repressor of floweringin *A. thaliana* that is regulated by vernalization. A period of 7 days of cold treatment at 4°C induced *CiMFL* full down-regulation under a long-day photoperiod, but not under a short-day photoperiod. This down-regulation resulted to be maintained after the cold treatment ended. Furthermore, together with a decrease in *CiMFL* transcripts, cold conditions induced changes in the morphology of the shoot apical meristem. Long-day photoperiod itself was not able to induce flowering.

Flowering begins in May to June: the azure-blue flowers open early in the morning and, under optimal light and temperature conditions, anthesis is complete before 10 a.m. The flowers usually do not stay open after the morning, unless under conditions of moderate temperature, shadow and high humidity. Seeds mature at 50 to 60 days after flowering and are collected in July to August. As the seed heads on the plant do not mature uniformly, particularly under conditions of high temperature and aridity, some shattering can occur when the seeds are collected. This phenomenon could be promoted by the birds that tend to feed on the mature flower heads. Nonetheless, the seeds need to be harvested at a suitable developmental stage because germination could be poor if they are not completely mature. In addition to its genetic value, the intrinsic and extrinsic properties of the seed depend upon the crop's growing conditions as well as on harvest techniques; overall, these facets determine the commercial quality and field performance of the seed. Harvesting is preferably carried out early in the morning by cutting the seed stalks at their base just before the seeds have dried out completely; in fact, the seeds may fall off the stalk and be lost if they are allowed to fully dry on the plant while still in the field. When the seed stalks are dry, the seed is trashed using adequate equipment and care to allow for optimal genetic purity and physical properties. The seed yield depends on the plant density and architecture, averaging 10 to 15 g per plant, hence a total production of 05-0.7 t/ha can be expected.

Depending on the cultivar, the 1,000 seed weight is 1.4 to 1.7 g (600 to 700 seeds per gram).

Due to their angular shape, commercial chicory seeds are often coated and pelletized with various materials (*e.g.*, cementite) to afford them with a more spherical shape. This facilitates the use of planting machines, pesticide application against seed and soil-borne pathogens, and priming processing to improve sowing performance. Because leaf chicory F_1 hybrids have entered the seed market, special attention should be given to the multiplication of parental lines and to F_1 seed production. Well established *in vitro* techniques or less sophisticated *in vivo* procedures based on the ability of sliced roots to produce numerous clones, may allow for the maintenance and adequate multiplication of chicory parental lines without any risk of genetic contamination.

Since *C. intybus* is a biennial species, plants do not enter the reproductive stage without vernalization [78-79-80]. This is why both Witloof and Radicchio, the commercial product of which is a head of tighten leaves, are commonly grown as annual plants but must be sown directly in the field in late spring or in summer to produce a head between October and February; alternatively, as a response of the forcing procedure, planting can occur sometime during the winter months according to the producer's choice. Regardless, anticipated or early flowering (bolting) caused by early spring sowing or late sowing and the sudden drop in temperature has to be avoided. This is because of the net loss in commercial production in addition to the shedding of unwanted seeds, resulting in weed chicory in the field of the following year. This situation poses two problems: a) the need for non-bolting varieties able to render the crop at least partially independent of the temperature during the first stage of development and b) the need for technical procedures able to permit selection and seed production of the selected plants in the same growing season, early enough to proceed according to an annual cycle.

With regard to the first issue, breeding programs could help in the development of new varieties resistant to early flowering.. One study on root chicory suggested an absolute vernalization requirement because bolting never occurred in control plants not subjected to low temperatures but maintained in long days in the field. However, because the effects of environmental parameters other than low temperature were not investigated, the same study could not definitively conclude that root chicory is an absolute cold-requiring plant [83].

Thus, the exact requirements for floral initiation in terms of the duration and intensity of an effective vernalization treatment and the developmental stage during which seeds and/or plants are sensitive to cold treatments remain to be established. Some research suggests that an essential factor for chicory flowering is the post-vernalization requirement for long days [81], but it is not known whether temperature is as determining as long days during this phase. It is quite evident that better knowledge of the genetics governing flower induction and vernalization requirements in chicory is fundamental for developing varieties with predictable flowering behavior.

In recent years, useful information on the genetic regulation of flowering time has been derived from the model plant *A. thaliana* [84]. Indeed, different classes of commercial earliness have been selected within each of the various types of Radicchio currently grown and to a certain extent, they show correspondence with as many classes of earliness in flowering time. Furthermore, it is noteworthy that the timing and rhythm of selection cycles might be quite different if one refers to the field head-producing types of Radicchio or to the forced chicories such as Witloof and Radicchio Late Red of Treviso. For these two chicories, selection mainly occurs after forcing on the basis of observations made on the commercial product. The selected plants can be transplanted in pots and later, after a period of acclimation, directly in the field or under an isolation cage.

The second issue thus mainly relates to the other four types of Radicchio and what follows herein refers to the procedure generally applied in north-eastern Italy, their typical area of production. Three aspects need to be considered: a) the moment of selection, b) the conservation of the taproot, and c) the timing of seed stalk development and seed production. The moment of selection strictly depends on the earliness of the material. In any case, the selection procedure, mainly based on field observations carried out during autumn or winter months and that usually suggest the destruction of the head, must take into account the necessity of preserving the roots of the selected plants until the following spring, the time when they will be transplanted in the field to obtain seed. Once selection has been completed, the selected plants are dug up, and the leaves are removed with a sharp cut at the base of the head, paying close attention to preserve the upper portion of the root where axillary buds are present in great number. Once washed and treated with fungicide, the taproots are transplanted into pots and maintained under plastic tunnels where

temperature should not reach below 0-2°C. During storage, the roots can start to bud, and it may be necessary to thin out the young sprouts. The roots of early material, selected in September or October, are often cold stored, both before and/or after transplanting, to inhibit sprouting. At the beginning of January, the preserved roots are transplanted in a protected environment where, with the aim of accelerating the flowering process, the temperature should be raised to 18°C and artificial light (150-250 lux at the vegetation level) should be added from 1-2 hours before sunset until midnight. The plant density adopted is 2 plants per square meter, which should result in production of approximately 40 g of viable seed per plant. During this phase, periodic defoliation of the basal portion of the seed stalk might be necessary to facilitate air circulation, and careful disease and insect control should be carried out. As the seed stalk grows, the plants should be clipped to 100-120 cm to promote earliness and simultaneous flowering. At the opening of the first flowers, pollinating agents (bees, flies or, more recently, bumblebees, according to the available volume) must be introduced under the isolation cage while drastically reducing and attentively controlling the use of any chemical spraying. The flowering period may last two to three months, with a peak between the third and the sixth weeks. This means that, independently of any attempt to anticipate flowering, harvesting cannot occur later than the end of mid June to enable the thrashing of the plants and preparation of the seed in time for the mid-July to mid-August sowing.

Concluding Remarks

In conclusion, it is worth adding some observations related to the significance that chicory has to horticulture within a European framework. The Witloof and Radicchio cultivar groups overall by far represent the largest quantity of chicory produced in Europe and elsewhere. The Radicchio group presents an almost astonishing phenotypic and genetic variability, from which selection is still able to generate new varieties and new commercially valuable forms. From a breeding point of view, Radicchio may represent a variable and rich germplasm resource with regard to useful morphological and/or physiological variants. Nevertheless, useful genes controlling some main quality traits or resistance to the most important pathogens are at very low frequency. In fact, farmer selection has traditionally paid attention to morphological characters, on which the market value mainly relies, whereas little attention has been given to other important characters, such as biotic or abiotic stress resistance, bitterness and post-harvest rotting. In both cases, useful genes may be identified, isolated and utilized through traditional breeding schemes or marker-assisted breeding methods through the adequate use of now-available molecular tools. From this point of view, the available numberless farmer populations of Radicchio represent an invaluable reserve of genetic resources that must be adequately studied, analyzed, classified, and compared to exploit them in breeding programs.

Indeed, *C. endivia* and *C. intybus* have the same number of chromosomes and are completely interfertile, and the same degree of sexual compatibility exists among different cultivars groups within each species and between these and the wild populations of both species. The complex *endivia-intybus* might thus be considered, from the breeder's point of view, as a unique genetic pool, the access to which is not limited by sexual incompatibility. As opposed to the underlying the possibility of transferring useful genes across species, cultivars groups, or cultivated and wild types through traditional methods, *i.e.*, selection of superior genotypes from segregating populations, this observation, although trivial, may suggest the feasibility of breeding programs aimed at creating horticultural novelties that might open new commercial perspectives or enlarge the existing ones. This possibility needs to be entirely explored by possibly integrating traditional procedures based on test crosses with more sophisticated molecular approaches.

References

[1]. Street RA, Sidana J, Prinsloo G (2013) *Cichorium intybus*: Traditional Uses, Phytochemistry, Pharmacology, and Toxicology. eCAM. DOI:10.1155/2013/579319.

[2]. Lucchin M, Varotto S, Barcaccia G, Parrini P (2008) Chicory and Endive. In Handbook of Plant Breeding, Vegetables I: Asteraceae, Brassicaceae, Chenopodicaceae. Springer Science. New York. p. 1-46. DOI:10.1007/978-0-387-30443-4_1.

[3]. Munoz CLM (2004) Spanish medicinal Plants: *Cichorium intybus* L. Boletin de la Real Sociedad Espanola de Historia Natura. 99.2: 41-47.

[4]. Plmuier W (1972) Chicory improvement. Revueder Agriculture. 4: 567-5.

[5]. FAOSTAT database collections. Food and Agriculture Organization of the United Nations, Rome. http://faostat3.fao.org/home/E. Accessed 06 June 2014.

[6]. Galla G, Ghedina A, Tiozzo CS, Barcaccia G (2016) Toward a first high-quality genome draft for marker-assisted breeding in leaf chicory, Radicchio (*Cichorium intybus* L.). In Plant Genomics, Dr. Ibrokhim Y. Abdurakhmonov (Ed.), InTech. DOI:10.5772/61747.

[7]. Bremer K and Anderberg A (1994) Asteraceae: cladistics and classification. Timber Press, Portland.

[8]. Tutin TG, Heywood VH, Burges NA, Moore DA, Valentine DH, Walters SM, Webb DA (1976) Flora Europea, vol. 4.

[9]. Pignatti S (1982) Flora d'Italia. 3 vols. Edagricole, Bologna.

[10]. Berdaff U (1985) Die Gattung Cichorium (Compositae), ihre Merkmale und ihre Arten, Georg-August-Universitat, Gottingen.

[11]. Kiers AM (2000) Endive, chicory and their wild relatives, a sistematic and phylogenetic study of Chicorium (Asteraceae)-Universiteit Leiden, Leiden.

[12]. Wagenitz G and Bedarff U (1989) Taxonomic notes on some species of the genus Cichorium (Compositae-Lactuceae). Edinburgh, University Press; 11-21.

[13]. Conti F, Abbate G, Alessandrini A, Blasi C (2005) An annotated checklist of the Italian vascular flora; Palombi Editori: Roma, Italy.

[14]. Vermeulen A, Desprez B, Lancelin D, Bannerot H (1994) Relationships among Cichorium species and related genera as determined by analysis of mitochondrial RFLPs. Theor. Appl. Genet. 88: 159-166. DOI:10.1007/BF00225892.

[15]. Gemeinholzer B and Bachmann K (2005) Examining morphological and molecular diagnostic character states of *Cichorium intybus* L. (Asteraceae) and *C. spinosum* L. Plant Syst. Evol. 253(1-4): 105-123. DOI:10.1007/s00606-004-0272-6.

[16]. Kiers AM, Mes TH, Van Der Meijden R, Bachmann K (1999) Morphologically defined Cichorium (Asteraceae) species reflect lineages based on chloroplast and nuclear (ITS) DNA data. Syst. Bot. 24: 645-659. DOI:10.2307/2419648.

[17]. Kiers AM, Mes TH, Van Der Meijden R, Bachmann K (2000) A search for diagnostic AFLP markers in Cichorium species with emphasis on Endive and Chicory cultivar groups. Genome. 43(3): 470-6. DOI:10.1139/g00-024

[18]. GRIN Database, Available online https://npgsweb.ars-grin.gov/gringlobal/search.aspx (Accessed 02 July 2014).

[19]. Mansfeld's World Database of Agriculture and Horticultural Crops Database Available online http://mansfeld.ipk-gatersleben.de/apex/f?p=185:3 (accessed 02 July 2014).

[20]. Raulier P, Maudoux O, Notté C, Draye X, Bertin P (2015) Exploration of genetic diversity within *Cichorium endivia* and *Cichorium intybus* with focus on the gene pool of industrial chicory. Genet. Resour. Crop. Ev. 1-17. DOI 10.1007/s10722-015-0244-4.

[21]. Van Stallen N, Noten V, Neefs V, De Proft M (2001) The phylogenetic relationship between different *Cichorium intybus* cultivars and cultivar groups, as revealed by RAPDs. Plant Breed. 120: 425-428. DOI:10.1046/j.1439-0523.2001.00633.x.

[22]. Koch G and Jung C (1997) Phylogenetic relationships of industrial Chicory varieties revealed by RAPDs and AFLPs. Agronomie. 17: 323-333. https://hal.archives-ouvertes.fr/hal-00885847/document

[23]. Van Stallen N, Vandenbussche B, Verdoodt V, De Proft M (2003) Construction of a genetic linkage map for witloof (*Chicorium intybus* L. var. *foliosum* Hegi). Plant Breed. 122: 251-525. DOI:10.1111/j.1439-0523.2003.00869.x.

[24]. De Proft MP, Van Stallen N, Veerle N (2003) Introduction: History of chicory breeding. Proc. of the Eucarpia meeting on Leafy Vegetables Genetics and Breeding, March 19-21, 2003, Noordwijkerhout, The Netherland. pp. 83-90.

[25]. Van Stallen NB, Vandenbussche B, Londers E, Noten V, De Proft M (2005) QTL analysis of production and taste characteristics of Chicory (*Cichorium intybus* var. *foliosum*). Plant Breed. 124: 59-62. DOI:10.1111/j.1439-0523.2004.01043.x.

[26]. Koukounaras A (2014) Yield and Quality Parameters of two Radicchio (*Cichorium intubys* L.) Cultivars as Affected by Growth Season. Eur. J. Hortic. Sci. 79(5) : 283-287. ISSN 1611-4426

[27]. Bianchedi A (1961) I radicchi di Treviso: storia, coltivazione, forzatura e commercio. Ital Agric. 98: 37-51.

[28]. Eenink AH (1981) Compatibility and incompatibility in witloof - Chicory (*Cichorium intybus* L.). 2. The incompatibility system. Euphytica. 30: 77-85. DOI:10.1007/BF00033662.

[29]. Varotto S, Pizzoli L, Lucchin M, Parrini P (1995) The incompatibility system in Italian red Chicory (*Cichorium intybus* L.). Plant Breed. 114: 535-538. DOI:10.1111/j.1439-0523.1995.tb00851.x.

[30]. Barcaccia G, Varotto S, Soattin M, Lucchin M, Parrini P (2003) Genetic and molecular studies of sporophytic self-incompatibility in *Cichorium intybus* L. Proc. of the Eucarpia meeting on Leafy Vegetables Genetics and Breeding, March 19-21, 2003, Noordwijkerhout, The Netherland, p. 154.

[31]. Kilian N, Gemeinholzer B, Lack HW (2009) Cichorieae. Systematics, evolution and biogeography of Compositae. IAPT, Vienna. 343-383.

[32]. Gonthier L, Blassiau C, Mörchen M, Cadalen T, Poiret M, Hendriks T, Quillet MC (2013) High-density genetic maps for loci involved in nuclear male sterility (*NMS1*) and sporophytic self-incompatibility (*S*-locus) in chicory (*Cichorium intybus* L., Asteraceae). Theor. Appl. Genet. 126(8) : 2103-2121.

[33]. Barcaccia G and Tiozzo CS (2012) New male sterile *Cichorium* spp. mutant, parts or derivatives, where male sterility is due to a recessive nuclear mutation linked to a polymorphic molecular marker, useful for producing F1 hybrids of *Cichorium* spp. EU PATENT No. WO2012163389-A1. DOI: 1013140/2.1.4749.5044.

[34]. Barcaccia G and Tiozzo CS (2014) New male sterile mutant of leaf chicory, including Radicchio, used to produce chicory plants and seeds with traits such as male sterility exhibiting cytological phenotype with shapeless, small and shrunken microspores in dehiscent anthers. USA PATENT No. US20140157448-A1. DOI: 10.13140/2.1.3176.6400.

[35]. Barcaccia G, Pallottini L, Soattin M, Lazzarin R, Parrini P, Lucchin M (2003) Genomic DNA fingerprints as a tool for identifying cultivated types of radicchio (*Cichorium intybus* L.) from Veneto, Italy. Plant Breed. 122: 178-183. DOI:10.1046/j.1439-0523.2003.00786.x.

[36]. Lucchin M, Barcaccia G, Lazzarin R, Parrini P (2003) E' in grado la selezione fenotipica di determinare differenziazione molecolare tra popolazioni in radicchio (*Cichorium intybus* L.)? Atti del Congresso della Società Orticola Italiana, Gruppo di Lavoro Miglioramento Genetico: Lo stato dell'arte nel miglioramento genetico delle principali specie ortoflorofrutticole d'interesse mediterraneo, Valenzano (Bari), Italy. 25-26 Giugno 2002, a cura di L. Ricciardi, pp. 387-396.

[37]. Ghedina A, Galla G, Cadalen T, Hilbert JL, Tiozzo CS, Barcaccia G (2015) A method for genotyping elite breeding stocks of leaf chicory (Cichorium intybus L.) by assaying mapped microsatellite marker loci. BMC Res Notes. 8(1) : 1-12. DOI:10.1186/s13104-015-1819-z.

[38]. Charlesworth D (2002) Plant sex determination and sex chromosomes. Heredity, 88(2): 94-101. DOI:10.1038/sj.hdy.6800016.

[39]. Chuck G (2010) Molecular mechanisms of sex determination in monoecious and dioecious plants. Adv. Bot. Res. 54: 53-83. DOI:10.1016/S0065-2296(10)54002-3.

[40]. Nettancourt D (2001) Incompatibility and incongruity in wild and cultivated plants (Vol. 3). Springer Science & Business Media, 2001. DOI: 10.1007/978-3-662-04502-2.

[41]. Allen AM and Hiscock SJ (2008) Evolution and phylogeny of self-incompatibility systems in angiosperms. In Self-Incompatibility in Flowering Plants, Springer Berlin Heidelberg. 73-101. DOI:10.1007/978-3-540-68486-2_4.

[42]. Hiscock SJ (2000) Self-incompatibility in *Senecio squalidus* L. (Asteraceae). Ann. Bot. 85 (supplement A), 181-190. DOI:10.1006/anbo.1999.1058, a.

[43]. Allen AM, Thorogood CJ, Hegarty MJ, Lexer C, Hiscock SJ (2011) Pollen–pistil interactions and self-incompatibility in the Asteraceae: new insights from studies of *Senecio squalidus* (Oxford ragwort). Ann. Bot. 108(4): 687-698. DOI:10.1093/aob/mcr147.

[44]. Hiscock SJ and McInnis SM (2003) Pollen recognition and rejection during the sporophytic self-incompatibility response: Brassica and beyond. Trends Plant Sci. 8(12): 606-613. DOI:10.1016/j.tplants.2003.10.007.

[45]. Castaño CI, Demeulemeester MAC, De Proft MP (1997) Incompatibility reactions and genotypic identity status of five commercial chicory (*Cichorium intybus* L.) hybrids. Sci Hortic-Amsterdam. 72(1): 1-9. DOI:10.1016/S0304-4238(97)00111-8.

[46]. Coppen's d'Eegkenbrugge G, Van Hergk JC, Dutilleul P (1989) A study of fructose yield components in Chicory. Plant Breed. 102: 296-301. DOI:10.1111/j.1439-0523.1989.tb01258.x.

[47]. Nasrallah JB (2000) Cell–cell signaling in the self-incompatibility response. Curr Opin Plant Biol. 3(5): 368-373. PMID: 11019803

[48]. Varotto S (1995) Personal communication.

[49]. McInnis SM, Costa LM, Gutiérrez-Marcos JF, Henderson CA, Hiscock SJ (2005) Isolation and characterization of a polymorphic stigma-specific class III peroxidase gene from *Senecio squalidus* L.(Asteraceae). Plant Mol. Biol. 57(5): 659-677. DOI:10.1007/s11103-005-1426-9

[50]. Kaul MLH (1988) Male Sterility in Higher Plants. SpringerVerlag: Berlin, 1988. DOl:10.1007/978-3-642-83139-3.

[51]. Horner HT and Palmer RG (1995) Mechanisms of genic male sterility. Crop Sci. 35(6): 1527-1535. DOI:10.2135/cropsci1995.0011183X003500060002x.

[52]. Budar F and Pelletier G (2001) Male sterility in plants: occurrence, determinism, significance and use. Comptes Rendus de l'Académie des Sciences-Series III-Sciences de la Vie. 324(6): 543-550. DOI:10.1016/S0764-4469(01)01324-5.

[53]. Wang DJ, Guo AG, Li DR, Tian JH, Huang F, Sun GL (2007) Cytological and molecular characterization of a novel monogenic dominant GMS in *Brassica napus* L. Plant Cell Rep. 26(5): 571-579. DOI:10.1007/s00299-006-0276-5.

[54]. Desprez BF, Delesalle L, Dhellemmes C, Desprez MF (1994) Genetics and breeding of industrial chicory. Comptes Rendus de l'Académie d'Agriculture de France. 80(7): 47-62. ISSN: 0989-6988

[55]. Cadalen T, Mörchen M, Blassiau C, Clabaut A, Scheer I, Hilbert JL *et al.*, Quillet MC (2010) Development of SSR markers and construction of a consensus genetic map for chicory (*Cichorium intybus* L.), Mol. Breeding. 25(4): 699-722. DOI:10.1007/s11032-009-9369-5.

[56]. Doré C and Varoquaux F (2006) Histoire et amélioration de cinquante plantes cultivées. Editions Quae, Inra, France. ISBN: 978-2-7380-1215-9.

[57]. De Neve C (2014) Male Sterility in Rice Pollen Development. Ghent University Library, Universiteit Gent. Faculteit Bio-ingenieurswetenschappen Vakgroep LA15.

[58]. Yamaguchi T, Lee DY, Miyao A, Hirochika H, An G, Hirano HY (2006) Functional diversification of the two C-class MADS box genes OSMADS3 and OSMADS58 in *Oryza sativa.* Plant Cell. 18(1): 15-28. DOI:10.1105/tpc.105.037200.

[59]. Collani S (2012) Reproductive barriers in crop plants: understanding the genetic and molecular bases of self-incompatibility in olive and male-sterility in chicory. PhD Thesis, University of Padova, Italy.

[60]. Chu MG, Li S, Wang S, Zheng A, Deng Q, Ding L, Zhang J, Zhang M, He M, Liu H, Zhu J, Wang L, Li P (2011) Fine mapping of a male sterility gene, *vr1*, on chromosome 4 in rice. Mol. Breed. 28: 181-187. DOI:10.1007/s11032-010-9472-7.

[61]. Reynaerts A, Van de Wiele H, de Sutter G, Janssens J (1993) Engineered genes for fertility control and their application in hybrid seed production. Sci. Hort. 55: 125-139. DOI:10.1016/0304-4238(93)90028-O.

[62]. Rambaud C, Dubois J, Vasseur J (1993) Male-sterile chicory cybrids obtained by intergeneric protoplast fusion. Theor. Appl. Gen. 87(3): 347-352. DOI:10.1007/BF01184922.

[63]. Rambaud C, Bellamy A, Dubreucq A, Bourquin JC, Vasseur J (1997) Molecular analysis of the fourth progeny of plants derived from a cytoplasmic male sterile chicory cybrid. Plant Breed. 116(5): 481-486. DOI:10.1111/j.1439-0523.1997.tb01035.x.

[64]. Dubreucq A, Berthe B, Asset JF, Boulidard L, Budar F, Vasseur J, Rambaud C (1999) Analyses of mitochondrial DNA structure and expression in three cytoplasmic male-sterile

chicories originating from somatic hybridisation between fertile chicory and CMS sunflower protoplasts. Theor. Appl. Gen. 99(7-8): 1094-1105. DOI:10.1007/s001220051314

[65]. Varotto S, Nenz E, Lucchin M, Parrini P (2001) Production of asymmetric somatic hybrid plants between *Cichorium intybus* and *Helianthus annuus.* Theor. Appl. Gen. 102: 6/7, pp. 950-956. DOI:10.1007/s001220000485.

[66]. Chen J, Hu J, Vick BA, Jan CC (2006) Molecular mapping of a nuclear male-sterility gene in sunflower (*Helianthus annuus* L.) using TRAP and SSR markers. Theor. Appl. Gen. 113(1): 122-127. DOI:10.1007/s00122-006-0278-2.

[67]. Jan CC and Vick BA (2007) Inheritance and allelic relationships of fertility restoration genes for seven new sources of male-sterile cytoplasm in sunflower. Plant Breed. 126(2): 213-217. DOI:10.1111/j.1439-0523.2007.01350.x.

[68]. Yue B, Vick BA, Cai X, Hu J (2010) Genetic mapping for the *Rf1* (fertility restoration) gene in sunflower (*Helianthus annuus* L.) by SSR and TRAP markers. Plant Breed. 129(1): 24-28. DOI:10.1111/j.1439-0523.2009.01661.x.

[69]. Habarugira I, Hendriks T, Quillet MC, Hilbert JL, Rambaud C (2015) Effects of Nuclear Genomes on Anther Development in Cytoplasmic Male Sterile Chicories (*Cichorium intybus* L.): Morphological Analysis. The Scientific World Journal. DOI:10.1155/2015/529521.

[70]. Bellamy A, Mathieu C, Vedel F, Bannerot H (1995) Cytoplasmic DNA and nuclear rDNA restriction fragment length polymorphisms in commercial witloof chicories. Theor. Appl. Genet. 91: 505-509. DOI:10.1007/BF00222980.

[71]. Van Stallen N, Noten V, Demeulemeester M, De Proft M (2000) Identification of commercial Chicory cultivars for hydroponic forcing and their phylogenetic relationships revealed by random amplified polymorphic DNAs and amplified fragment length polymorphisms. Plant Breed. 119: 265-270. DOI:10.1046/j.1439-0523.2000.00499.x.

[72]. Bellamy A, Vedel F, Bannerot H (1996) Varietal identification in *Cichorium intybus* L. and determination of genetic purity of F1 hybrid seed samples, based on RAPD markers. Plant Breed. 115: 128-32. DOI:10.1111/j.1439-0523.1996.tb00887.x.

[73]. Hodgins KA, Lai Z, Oliveira LO, Still DW, Scascitelli M, Barker MS *et al.* and Burke JM (2014) Genomics of Compositae crops: reference transcriptome assemblies and evidence of hybridization with wild relatives. Mol. Ecol. Resour. 14(1): 166-177. DOI:0.1111/1755-0998.12163.

[74]. Kiær LP, Felber F, Flavell A, Guadagnuolo R, Guiatti D, Hauser TP, Van De Wiel C (2009) Spontaneous gene flow and population structure in wild and cultivated chicory, *Cichorium intybus* L. Genet. Resour. Crop. Ev. 56(3): 405-419. DOI: 10.1007/s10722-008-9375-1.

[75]. De Simone M, Morgante M, Lucchin M, Parrini P, Marocco AA (1997) first linkage map of *Cichorium intybus* L. using a one-way pseudo-testcross and PCR-derived markers. Mol. Breed. 3: 415-425. DOI:10.1023/A:1009616801404.

[76]. Muys C, Thienpont CN, Dauchot N, Maudoux O, Draye X, Van Cutsem P (2013) Integration of AFLPs, SSRs and SNPs markers into a new genetic map of industrial chicory (*Cichorium intybus* L. var. *sativum*). Plant Breed. 113(1): 130-137. DOI:10.1111/pbr.12113.

[77]. Delporte M, Legrand G, Hilbert JL, Gagneul D (2015) Selection and validation of reference genes for quantitative real-time PCR analysis of gene expression in *Cichorium intybus*. Front Plant Sci. 18(6): 651. DOI:10.3389/fpls.2015.00651

[78]. Carazzone C, Mascherpa D, Gazzani G, Papetti A (2013) Identification of phenolic constituents in red chicory salads (*Cichorium intybus*) by high-performance liquid chromatography with diode array detection and electrospray ionisation tandem mass spectrometry. Food Chem. 138(2): 1062-1071. DOI:10.1016/j.foodchem.2012.11.060.

[79]. Pushparaj PN, Low HK, Manikandan J, Tan BKH, Tan CH (2007) Anti-diabetic effects of *Cichorium intybus* in streptozotocin-induced diabetic rats. J. Ethnopharmacol. 111(2): 430-434. DOI:10.1016/j.jep.2006.11.028.

[80]. Tousch D, Lajoix AD, Hosy E, Azay-Milhau J, Ferrare K, Jahannault C *et al*. Petit P (2008) Chicoric acid, a new compound able to enhance insulin release and glucose uptake. Biochem. Biophys. Res. Commun. Bio. 377(1): 131-135. DOI:10.1016/j.bbrc.2008.09.088.

[81]. Heimler D, Isolani L, Vignolini P, Romani A (2009) Polyphenol content and antiradical activity of *Cichorium intybus* L. from biodynamic and conventional farming. Food Chem. 114(3): 765-770. DOI:10.1016/j.foodchem.2008.10.010.

[82]. KEGG: Available online: http://www.genome.jp/kegg/ (accessed 28.08.2015).

[83]. Paulet P (1985) *Cichorium intybus* and *C. endiva*. In *CRC Handbook of Flowering*, Editor A.H. Halevy, CRC Press Inc., Boca Raton, FL. Volume 2, pp. 265-270.

[84]. Van Stallen N, Demeulemeester MAC, De Proft M (1999) An in vitro model for the study of hydroponic forcing in Chicory (*Cichorium intybus* L.). Plant Cell Tiss. Org. 55: 125-131. DOI:10.1023/A:1006209207540.

[85]. Gianquinto G (1997) Morphological and physiological aspects of phase transition in radicchio (*Cichorium intybus* L. var. s*ilvestre* Bisch.): influence of daylength and its interaction with low temperature. Sci Hortic-Amsterdam. 71(1): 13-26. DOI:10.1016/S0304-4238(97)00058-7

[86]. Wiebe HJ (1997) Ursachen für die generative Entwicklung von Radicchio (*Cichorium intybus* var *foliosum*). Gartenbauwissenschaft. 62: 72-77.

[87]. Demeulemeester MAC and De Proft MP (2000) Use of chicory root explants in vitro as a model for the induction of different stem types. J. Plant Physiol. 156: 413-418. DOI:10.1016/S0176-1617(00)80082-6.

[88]. Dielen V, Notté C, Lutts S, Debavelaere V, Van Herck JC, Kinet JM (2005) Bolting control by low temperature in root chicory (*Cichorium intybus* var *sativum*). Field Crop Res. 94: 76-85. DOI:10.1016/j.fcr.2004.11.005.

[89]. Amasino RM (2005) Vernalization and flowering time. Curr. Opin. Biotechnol. 16: 154-8. DOI:10.1016/j.copbio.2005.02.004.

Chapter II

A method for genotyping elite breeding stocks of leaf chicory (*Cichorium intybus* L.) by assaying mapped microsatellite marker loci

Published: Ghedina A, Galla G, Cadalen T, Hilbert JL, Tiozzo CS, Barcaccia G (2015) A method for genotyping elite breeding stocks of leaf chicory (*Cichorium intybus* L.) by assaying mapped microsatellite marker loci. BMC Res. Notes. 8(1): 1-12. DOI:10.1186/s13104-015-1819-z.

Abstract

Leaf chicory (*Cichorium intybus* subsp. *intybus* var. *foliosum* L.) is a diploid plant species (2n=18) of the Asteraceae family. The term "chicory" specifies at least two types of cultivated plants: a leafy vegetable, which is highly differentiated with respect to several cultural types, and a root crop, whose current industrial utilization primarily addresses the extraction of inulin or the production of a coffee substitute. The populations grown are generally represented by local varieties (*i.e.,* landraces) with high variation and adaptation to the natural and anthropological environment where they originated, and have been yearly selected and multiplied by farmers. Currently, molecular genetics and biotechnology are widely utilized in marker-assisted breeding programs in this species. In particular, molecular markers are becoming essential tools for developing parental lines with traits of interest and for assessing the specific combining ability of these lines to breed F1 hybrids.

The present research deals with the implementation of an efficient method for genotyping elite breeding stocks developed from old landraces of leaf chicory, Radicchio of Chioggia, which are locally dominant in the Veneto region, using 27 microsatellite (SSR) marker loci scattered throughout the linkage groups. Information on the genetic diversity across molecular markers and plant accessions was successfully assessed along with descriptive statistics over all marker loci and inbred lines. Our overall data support an efficient method for assessing a multi-locus genotype of plant individuals and lineages that is useful for the selection of new varieties and the certification of local products derived from Radicchio of Chioggia. This method proved to be useful for assessing the observed degree of homozygosity of the inbred lines as a measure of their genetic stability; plus it allowed an estimate of the specific combining ability (SCA) between maternal and paternal inbred 54 lines on the basis of their genetic diversity and the predicted degree of heterozygosity of their F1 hybrids. This information could be exploited for planning crosses and predicting plant vigor traits (*i.e.*, heterosis) of experimental F1 hybrids on the basis of the genetic distance and allelic divergence between parental inbred lines. Knowing the parental genotypes would allow us not only to protect newly registered varieties but also to assess the genetic purity and identity of the seed stocks of commercial F1 hybrids, and to certificate the origin of their food derivatives.

Key Words

Radicchio/MAB/inbred lines/SSR markers/genotyping

Introduction

Cichorium (*Cichorium intybus* subsp. *intybus* var. *foliosum* L.) comprises diploid plant species (2*n*=18) belonging to the Asteraceae family, subfamily Cichoriodeae, tribe Lactuceae or Cichorieae. These species are biennial or, in the wild, perennial species [1]. They are naturally allogamous due to an efficient sporophytic self-incompatibility system. In addition, outcrossing is promoted by a floral morpho-phenology unfavorable to selfing in the absence of pollen donors (*i.e.*, proterandry, wherein the anthers mature before the pistils) and a favorable competition of allo-pollen grains and tubes (*i.e.*, pollen that is genetically diverse from that produced by the seed parents, usually called auto-pollen) [2].

Long appreciated as medical plants by the ancient Greeks and Romans, *Cichorium* spp. are currently among the most important cultivated vegetable crops. They are generally used as components in fresh salads or, more rarely, cooked according to local traditions and alimentary habits [1].

Lacking comprehensive, homogeneous, sufficiently detailed, and univocal data on horticultural production and trade, it is difficult to give reliable figures on the diffusion and economic importance of this culture in Europe, where it is predominantly grown [1]. In recent statistics on the European market [3], chicory is often included under the general heading “salads” or considered together with lettuce, which is by far the most important leafy vegetable on both a European and world-wide scale. On the basis of accessible data [3], however, it is possible to determine that chicory is produced almost exclusively by Belgium, France, Italy, and the Netherlands [3]. Although chicory does not contribute greatly to each country’s total agricultural income, in the north eastern regions of Italy, it accounts for 87% of the national acreage and 84% of the national production of the red or variegated chicory known as “Radicchio”, which traditionally includes all the cultivar groups with leaf commercial products. This particular type of chicory is now receiving greater attention in Europe and the USA, where its cultivation originated several years ago, and is becoming increasingly subjected to evaluation because its red or variegated leaves are appreciated as a component of ready-to-eat salads [1].

The materials grown are generally represented by actively cultivated local populations (*i.e.*, landraces) with high variation and adaptation to the natural and anthropological environment where they originated, and have been yearly selected and multiplied by

farmers [1]. These populations are maintained by farmers through phenotypical selection based on their own criteria and occasionally on the exploitation of controlled hybridizations of different types to obtain recombinant genotypes that exhibit superior agronomic and commercial traits [4]. However, conventional plant breeding methods for hybridizing and selecting plants on the basis of observed phenotypes are not the only methods used by plant breeders. Currently, molecular genetics and biotechnology are widely utilized in breeding programs of the vast majority of crop plant species. Indeed, molecular markers are nowadays essential tools to select pure or inbred lines with qualitative traits of agricultural interest by marker-assisted selection (MAS) and also to predict the specific combining ability of parental lines to breed F1 hybrid varieties in marker-assisted breeding (MAB) schemes. In leaf chicory, molecular markers can find utility for assessing the degree of homozygosity of parental inbred lines, as a measure of their genetic stability, and also for predicting the degree of heterozygosity of their F1 hybrids, as an estimate of the specific combining ability on the basis of the genetic diversity between maternal and paternal inbred lines. Here we describe a method for genotyping elite breeding stocks – inbred lines – using microsatellites, or SSR (Simple Sequence Repeat) markers. Among the different PCR derived molecular systems, SSR markers are suitable for population genetics studies and marker-assisted selection programs because they have several desirable features, such as a high level of reproducibility, co-dominant heredity, and no need for high throughput technology [5]. Moreover, they provide loci for simple and accurate individual typing in any species and display a high level of polymorphism and widespread distribution in the genomes [6]. Differently from other single-locus marker systems, like for instance SNP (Single Nucleotide Polymorphisms) markers, microsatellite-based markers require a much less preliminary genomic information and bioinformatic characterization for their exploitation in a given species [7].

A large-scale application of molecular marker techniques, including amplified fragment length polymorphism (AFLP) and random amplified polymorphic DNA (RAPD), were used to construct the first genetic linkage maps of *C. intybus* [8-9]. In 2010, a new genetic linkage map for *C. intybus* was constructed by using SSR markers [10]. This consensus genetic map, which includes 9 homologous linkage groups, was obtained after the

integration and ordination of the molecular marker data deriving from one witloof chicory and two industrial chicory progenies.

The aim of our study was to develop a method for the genetic characterization of elite inbred lines of the “Red of Chioggia” chicory using mapped SSR markers with a particular emphasis on the assessment of the genetic stability within (*i.e.*, observed degree of homozygosity) and genetic diversity between paternal and maternal lines (*i.e.*, expected degree of heterozygosity of their F1 hybrids). Information derived from the application of this method should then be exploited for planning crosses and predicting plant vigor traits (*i.e.*, heterosis) of experimental F1 hybrids of leaf chicory on the basis of the genetic distance and allelic divergence between parental inbred lines. Knowing the parental genotypes would allow us not only to protect newly registered varieties but also to assess the genetic purity and identity of the seed stocks of commercial F1 hybrids, and to certificate the origin of their food derivatives.

Basic genetic variation and differentiation statistics computed for single locus, across linkage groups and plant accessions are presented and discussed. Overall data support an efficient method for assessing a multi-locus genotype of plant individuals and lineages, which can also be combined with pedigree notes on a panel of morpho-phenological traits for breeding F1 hybrid varieties in the Radicchio of Chioggia biotype.

Materials and Methods

Plant materials and DNA isolation

Plant materials of the "Red of Chioggia" biotype, belonging to *C. intybus* subsp. intybus var. *foliosum* L. (**Table 1**), were developed and provided by T&T Produce (Sant'Anna di Chioggia, Venice, Italy).

Table 1. Information on the plant materials, including the inbred line ID, the number of individuals assayed per population, the inbred level reached per each line and the cycle of the variety each line derives from, expressed in days after transplanting. The number of generations of selfing (S) is reported for pollen donors, whereas full-sibling (FS), backcrossing (BC), pair-wise crossing (F) between inbred lines or inter-crossing between selfed individuals (IS) refers to seed parents. In addition, male inbred lines were multiplied *in vivo* by seeds (*i.e.*, inbreeding) and female inbred lines were propagated *in vitro* by cuttings (*i.e.*, cloning).

Accession ID	No. individuals	Population type	Varietal cycle (d)
P04	8	S4	55
P31	8	S4	55
P33	8	S4	55
P11	8	S5	55
S02	8	S4	65
S03	8	S4	65
S31IS3.2	6	IS3	65
S31IS3.4	9	IS3	65
S31S5.2	12	S5	65
S31S3	8	S3	65
S31S5.1	3	S5	65
Z22	8	S4	75
Z23	8	S4	75
Z31	8	S5	75
Z33	8	S4	75
Z34	8	S4	75
U22	8	S2	75
U24	8	S2	75
QC03	8	S5	90
QC31	8	S5	90
CS441	8	S6	110
CS501	8	S6	110
SC24	8	S5	120
SE802	8	S6	140
SE902	8	S6	140
SE111S6	8	S6	140
SEG111	1	S6	140
SE111S7	8	S7	140
SE412	8	S6	140
SE501	8	S6	140

13	1	S1BC1	65
11	1	F2	80
17	1	FS1	90
49	1	F2	90
20	1	F2	110
38	1	FS1	110
86	1	F2	110

Most of the inbred lines were represented by pollen donors spanning from S3 to S7 obtained by repeated selfing of single individuals chosen within each progeny at both genotype and phenotype levels (see upper part of **Table 1**).

Lines coded as IS3 were partial inbreds derived from intercrossing closely related S3 progeny plants. A few male-sterile seed parents were selected within F2 progenies obtained by selfing of F1 individuals, FS1 progenies produced by full-sibling and S1BC1 progenies generated by selfing of BC1 individuals (see lower part of **Table 1**). Concerning the strategy of samp0ling, the pollen parents were represented by 8 individuals per accession, with few exceptions, whereas the seed parents were propagated by in vitro culture and represented each by replicated individuals of clonal lines. All plant materials were bred in an experimental station at Chioggia (Venice, Italy) under controlled pollination conditions.

The genomic DNA was extracted from 100 mg of fresh leaves with the GenElute™ Plant Genomic DNA Miniprep Kit (Sigma-Aldrich) following the manufacturer's instructions. The quality of the DNA samples was assessed by electrophoresis on 1% (w/v) agarose gel stained with 1X SYBR® Safe™ DNA Gel Stain (Life Technologies) in Tris-Acetate-EDTA (TAE) running buffer. The yield and purity of the extracted genomic DNA samples were evaluated using a NanoDrop 2000c UV-Vis Spectrophotometer (Thermo Scientific). Following DNA quantification, all the DNA samples were diluted to a final concentration of 25 ng/µl to be used as template for PCR amplifications.

Amplification of SSR loci

A total of 27 SSR marker loci were selected among those mapped in the 9 basic linkage groups constructed for *C. intybus* [10]. In particular, three SSR loci were carefully chosen for each linkage group in order to select the best ones in terms of polymorphism information content (PIC) scores and also to be well scattered throughout the genetic map (**Figure 1**).

The amplification of microsatellites was performed by using a PCR multiplex assay and the detection of DNA fragments across marker loci was achieved using a 5' M13-tailed primer method [12] with some modifications. Only one dye-labeled M13 primer was used per PCR reaction in combination with any other M13-tailed forward primer [14].

The SSR motifs and primers used in this study are described in **Appendix 1.1.**

Amplification reactions were set in order to analyze two marker loci with the same fluorescent dye in each PCR experiment [14]. In general, the marker alleles produced for different target loci, assayed with the same fluorescent dye labels, were characterized by distinct amplicons deriving from PCRs of similar efficiency, as DNA markers were combined in order to obtain reproducible amplicons of different size ranges.

The reactions were performed in a total volume of 20 μl including 2X Platinum® Multiplex PCR Master Mix, 10X GC Enhancer (Applied Biosystems, Carlsbad CA), 0.25 μM of each tailed primer, 0.75 μM of each non-tailed primer, 0.5 μM of each labeled primer (Invitrogen, Carlsbad CA), 25 ng of DNA and distilled water. The amplification reactions were performed in a Gene Amp® PCR System 9700 thermal cycler (Applied Biosystems).

The microsatellite-containing regions were amplified using two different PCR cycles defined to maximize the amplification efficiency of the different primer sets according to their annealing temperatures, as reported in **Appendix 1.1,** as well as the fluorescent dye labels and the two-loci matching system to perform the multi-locus PCR reactions.

The cycle termed 'rad-multi 54' consisted of an amplification reaction performed under the following conditions: 95°C for 5 min followed by 5 cycles at 95°C for 30 sec, 58°C for 45 sec decreasing by 0.8°C every cycle, 68°C for 45 sec followed by 35 cycles at 95°C for 30 sec, 54°C for 45 sec, 68°C for 45 sec and a final extension of 60 min at 68°C. Similarly, the cycle 'rad-multi 56' consisted of the following conditions: 95°C for 5 min followed by 5 cycles at 95°C for 30 sec, 60°C for 45 sec decreasing by 0.8°C every cycle, 68°C for 45 sec followed by 35 cycles at 95°C for 30 sec, 56°C for 45 sec, 68°C for 45 sec and a final extension of 60 min at 68°C.

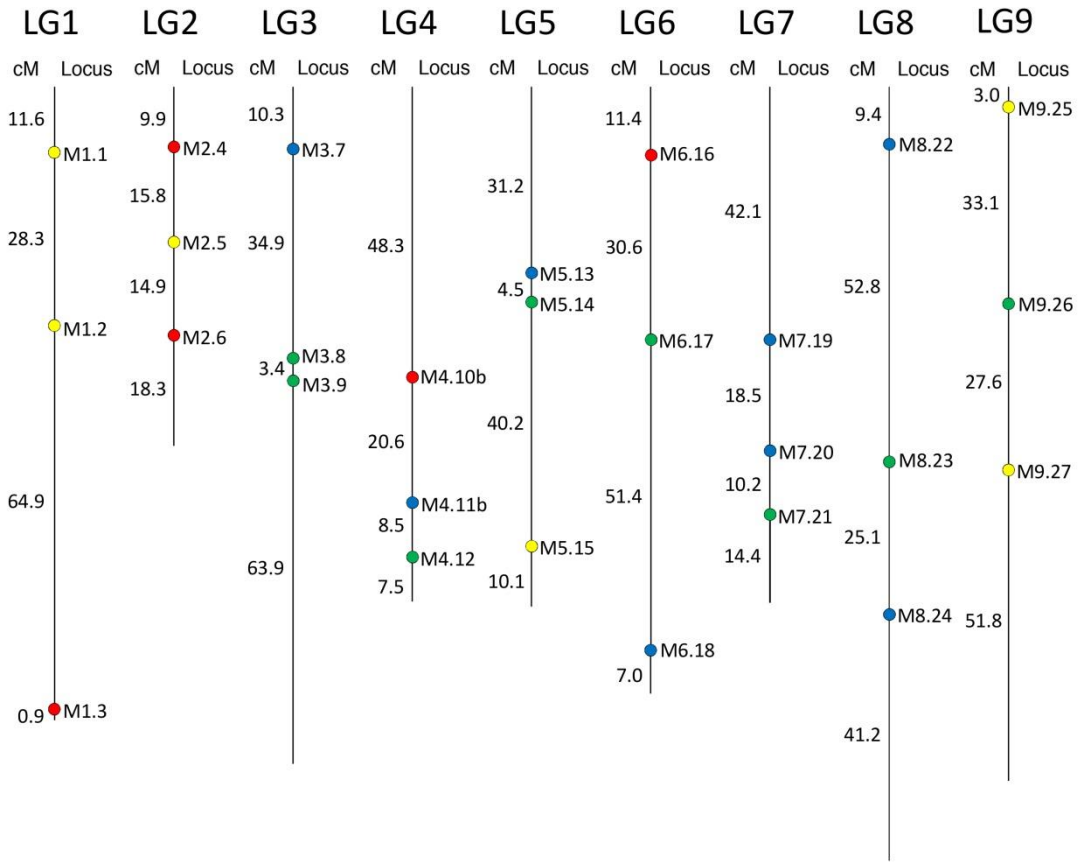


Figure 1. Consensus genetic linkage map of chicory (*C. intybus*), modified from Cadalen and coll. [10]. The colored dots indicate the positions of the selected marker loci throughout the nine basic linkage groups. Each color represents the fluorophore used to label microsatellite-containing amplicons related to each locus. The position of the marker loci and the length of each linkage group are also reported.

The quality of the PCR products was assessed by electrophoresis on a 2% (w/v) Agarose gel stained with 1X SYBR® Safe™ DNA Gel Stain (Life Technologies) using Tris-Acetate-EDTA (TAE) running buffer. The concentrations of the PCR products were estimated with KODAK 1D Image Analysis Software by comparing the intensity of the PCR products to that of a 1Kb Plus DNA Ladder (Life Technologies). Then, for each PCR reaction and fluorophore, approximately 20 ng of amplification products were pooled and prepared for capillary electrophoresis.

Analysis of the SSR loci

DNA fragment capillary electrophoresis was completed at BMR Genomics (Padova, Italy). Following electrophoresis, fragment analyses were performed with Peak Scanner^{TM} v. 1.0 (Life Technologies).

Owing to the relatively recent origin of the biotype "Red of Chioggia" (*i.e.*, years 1960-70's), belonging to the same botanical variety (*i.e.*, *Cichorium intybus* subsp. *intybus* var. *foliosum* L.), we assumed the absence of homoplasy and adopted an infinite alleles model, hence considering that marker alleles of the same size at a given locus had the same evolutionary history. Statistical analyses of the SSRs were performed with the PopGene software package v. 1.32 [15] for calculating allele frequencies: the alleles per locus (ne), Levene's [16] observed heterozygosity (Ho), the expected heterozygosity (He) and the average heterozygosity (Ha) were computed for each locus per line. PopGene, software was also used to estimate F-statistics [17]. Estimates of the heterozygosity within (Fis) and between (Fit) subpopulations were determined, as was the fixation index (Fst) according to Wright [18].

The phenotypic diversity of the marker allele profiles was estimated using Shannon's information index (I) as reported by Lewontin [19]. Gene flow (Nm) estimates among the subpopulations were derived from the fixation index as described by McDermott and McDonald [20].

The NJ dendrogram of all the accessions was based on Nei's method [21]. A bootstrap analysis was conducted with 1000 resampling replicates. The Dice's coefficient [22] was applied to calculate the proportion of genetic similarity (GS) in all the pair-wise comparisons of individuals. Values of the genetic similarity calculated with Rohlf's coefficient were used to conduct PCA, and the results are represented as centroids plotted according to the MGS (Mean Genetic Similarity) estimates.

All the relevant calculations and analyses were conducted using the appropriate routines in the NTSYS software package v. 2.21c [23]. The population structure of the inbred red Chioggia chicory lines was investigated using the model-based (Bayesian) clustering algorithm implemented in the STRUCTURE software [24], which groups individuals according to marker allele combination and distribution. All the simulations were executed assuming an admixture model with no a priori population information. The calculations

were performed with 300,000 iterations and 300,000 burn-ins under the assumption that the allele frequencies in the populations were correlated. Fifteen replicate runs were performed, with each run exploring a range spanning 1 to 35 K. The most likely value of K was estimated using ΔK, as reported in other studies [11]. The individuals with membership coefficients of $qi>0.8$ were assigned to specific groups, whereas the individuals with $qi<0.8$ were characterized identified as being admixed.

Results

PCR-based amplifications of the genomic DNA samples from all inbred lines were performed to assay 27 mapped loci (information on primer pairs is reported in **Appendix 1.1**).

Descriptive statistics over all the SSR loci, along with information on the genetic diversity found across the molecular markers and plant accessions, are reported in **Table 2** and **Table 3,** respectively. The mean number of observed marker alleles (na) in the SSR loci assayed was 5.9, varying from 3.3 in LG5 and LG7 to 9.3 in LG2 (**Table 2**).

Table 2. Descriptive statistics of the SSR marker loci: the sample size of individual enotypes (N), the frequency of the most common marker allele (pi), estimates of Shannon's information index of phenotypic diversity (I), the average number of observed alleles (na) and the effective number of alleles (ne) per locus, the observed heterozygosity (Ho), the expected heterozygosity computed using Levene (He), the average heterozygosity (Ha), Wright's inbreeding coefficients Fis and Fit, the fixation index (Fst), and gene flow (Nm).

		General statistics				H-statistics			F-statistics			Gene Flow
Locus	**N**	**p_i**	**I**	**n_a**	**n_e**	**H_o**	**H_e**	**H_a**	**F_{is}**	**F_{it}**	**F_{st}**	**Nm**
M1.1	400	0.423	0.741	3	2.009	0.165	0.503	0.117	0.017	0.827	0.824	0.054
M1.2	482	0.407	1.497	7	3.612	0.315	0.725	0.247	-0.055	0.640	0.658	0.130
M1.3	480	0.467	1.374	8	3.007	0.342	0.669	0.297	-0.154	0.471	0.542	0.212
LG1 Mean	**454**	**0.432**	**1.204**	**6.0**	**2.876**	**0.274**	**0.632**	**0.220**	**-0.064**	**0.646**	**0.675**	**0.132**
M2.4	482	0.317	1.667	9	4.703	0.291	0.789	0.263	-0.249	0.588	0.670	0.123
M2.5	484	0.401	1.370	8	3.357	0.252	0.704	0.250	-0.086	0.610	0.641	0.140
M2.6	486	0.438	1.792	11	4.061	0.251	0.755	0.244	-0.134	0.634	0.677	0.119
LG2 Mean	**484**	**0.386**	**1.610**	**9.3**	**4.040**	**0.265**	**0.749**	**0.252**	**-0.156**	**0.611**	**0.663**	**0.127**
M3.7	480	0.592	1.096	7	2.335	0.213	0.573	0.202	-0.194	0.590	0.657	0.131
M3.8	302	0.669	0.880	4	1.994	0.007	0.500	0.003	-0.067	0.995	0.996	0.001
M3.9	478	0.523	1.220	5	2.746	0.239	0.637	0.213	-0.320	0.580	0.682	0.117
LG3 Mean	**420**	**0.595**	**1.065**	**5.3**	**2.358**	**0.153**	**0.570**	**0.139**	**-0.193**	**0.722**	**0.778**	**0.083**
M4.10b	486	0.410	1.225	4	3.074	0.165	0.676	0.194	0.014	0.720	0.716	0.099
M4.11b	482	0.847	0.481	3	1.360	0.017	0.265	0.052	0.632	0.927	0.801	0.062
M4.12	480	0.506	1.661	10	3.357	0.208	0.704	0.191	-0.103	0.699	0.727	0.094
LG4 Mean	**483**	**0.587**	**1.122**	**5.7**	**2.597**	**0.130**	**0.548**	**0.146**	**0.181**	**0.782**	**0.748**	**0.085**
M5.13	476	0.517	1.129	4	2.684	0.202	0.629	0.184	-0.141	0.673	0.714	0.100
M5.14	474	0.884	0.359	2	1.258	0.114	0.206	0.102	-0.264	0.383	0.511	0.239
M5.15	478	0.699	0.754	4	1.796	0.092	0.444	0.074	-0.282	0.783	0.831	0.051
LG5 Mean	**476**	**0.700**	**0.748**	**3.3**	**1.913**	**0.136**	**0.426**	**0.120**	**-0.229**	**0.613**	**0.685**	**0.130**

M6.16	476	0.546	1.304	5	2.835	0.181	0.649	0.177	-0.145	0.685	0.725	0.095
M6.17	480	0.469	1.707	9	3.750	0.358	0.735	0.293	-0.269	0.500	0.606	0.163
M6.18	480	0.527	1.495	8	3.063	0.208	0.675	0.169	-0.183	0.703	0.749	0.084
LG6 Mean	**479**	**0.514**	**1.502**	**7.3**	**3.216**	**0.249**	**0.686**	**0.213**	**-0.199**	**0.629**	**0.693**	**0.114**
M7.19	480	0.554	0.996	3	2.460	0.479	0.595	0.346	-0.622	0.052	0.416	0.352
M7.20	484	0.531	0.739	4	2.023	0.236	0.507	0.241	-0.305	0.402	0.542	0.211
M7.21	474	0.878	0.390	3	1.275	0.076	0.216	0.099	-0.165	0.590	0.648	0.136
LG7 Mean	**479**	**0.654**	**0.708**	**3.3**	**1.919**	**0.264**	**0.439**	**0.228**	**-0.364**	**0.348**	**0.535**	**0.233**
M8.22	480	0.990	0.058	2	1.021	0.021	0.021	0.016	-0.069	-0.009	0.056	4.192
M8.23	482	0.512	1.501	9	3.143	0.261	0.683	0.210	-0.307	0.605	0.698	0.108
M8.24	478	0.398	1.664	8	4.144	0.243	0.760	0.195	-0.245	0.675	0.739	0.088
LG8 Mean	**480**	**0.633**	**1.074**	**6.3**	**2.769**	**0.175**	**0.488**	**0.140**	**-0.207**	**0.424**	**0.498**	**1.463**
M9.25	484	0.686	0.763	4	1.828	0.132	0.454	0.138	0.178	0.732	0.674	0.121
M9.26	478	0.490	1.354	6	3.101	0.088	0.679	0.176	0.459	0.860	0.741	0.087
M9.27	482	0.386	1.792	10	4.611	0.440	0.785	0.341	-0.327	0.425	0.567	0.191
LG9 Mean	**481**	**0.520**	**1.303**	**6.7**	**3.180**	**0.220**	**0.639**	**0.218**	**0.103**	**0.672**	**0.660**	**0.133**

The frequency of the most common marker allele (pi) proved to be low when the observed number of marker alleles was high and vice versa (for instance, in LG2 and LG5, where the average pi was 0.386 and 0.700, respectively). At the same time, both the expected heterozygosity (He) and Shannon's information index of phenotypic diversity (I) were estimated to be high when the observed number of marker alleles was high (for instance, in LG2, where the average pi was 0.749 and 1.610, respectively) (for additional statistics, see **Table 2**).

The observed homozygosity scores were high, as expected for inbred lines, with a mean estimate of 0.793 (st. dev.=0.120), and ranged from 0.521 to 0.993. The marker loci M3.8, M4.11b and M8.22, with observed heterozygosity values of 0.007, 0.017 and 0.021 (*i.e.*, homozygosity rates of 0.993, 0.983, 0.979), respectively, greatly contributed to this average homozygosity. Wright's inbreeding coefficients (F-statistics) for single marker loci were also computed (**Table 2**). The inbreeding coefficient calculated for individual accessions revealed a negative value, on average equal to Fis=−0.125, as shown in **Table** 2. This feature was shared by 22 of the 27 SSR marker loci investigated, and it was particularly evident for the marker locus M7.19, which scored a very low observed homozygosity of 0.521. Values of the Wright's fixation index, which were computed for each locus across linkage groups, are reported in **Table 2**. The average value was Fst=0.659. Estimates of gene flow (Nm) were also computed for each locus (**Table 2**). The calculated values were

slightly Nm>0 for the vast majority of the assayed marker loci, ranging from a minimum of 0.001 to a maximum of 0.352, with an average value equal to Nm=0.278.

Regarding the descriptive statistics over all the accessions, the number of polymorphic loci 175 among individuals within inbred lines varied from 5 (18.5%) to 23 (85.2%) out of the total of 27 marker loci (**Table 3**). Simple Matching coefficients scored values greater than 0.900 in all the accessions and were equal to 1.000 in eight. The observed homozygosity was on average high, with values greater than 0.800 in the majority of the inbred lines (ranging between 0.539 and 0.898). Conversely, the expected heterozygosity was on average law, varying from 0.098 to 0.433 (**Table 3**).

Principal coordinate analysis allowed for the definition of centroids for all the lines. The first two principal components explained 30.37% of the total genetic variation found within the analyzed lines. Specifically, the first and second components explained more than 18% and about 12% of the total genetic diversity, respectively.

Table 3. Descriptive statistics over all the accessions: the number of polymorphic loci (nPl), their frequency presented as percentage (%) of polymorphic loci on a total of 27 assayed, estimates of Shannon's information index of phenotypic diversity (I), the genetic similarity coefficient or Simple Matching coefficient (SM), the average number of observed alleles (na) and the effective number of alleles (ne), the observed heterozygosity (Ho), the expected heterozygosity computed using Levene (He), the average heterozygosity (Ha), Wright's inbreeding coefficients Fis and Fit and the fixation index (Fst).

	General statistics						H-statistics			F-statistics	
Line	**nPl**	**%**	**I**	**SM**	n_a	n_e	H_o	H_e	H_a	F_{is}	F_{st}
P04	11	40.74	0.240	0.968	1.407	1.301	0.144	0.177	0.186	0.189	0.693
P31	18	66.67	0.429	0.925	1.741	1.573	0.291	0.313	0.186	0.070	0.457
P33	16	59.26	0.379	0.927	1.615	1.472	0.260	0.280	0.193	0.073	0.513
P11	11	40.74	0.279	0.950	1.423	1.379	0.234	0.212	0.193	-0.102	0.632
S02	8	29.63	0.210	0.976	1.346	1.281	0.130	0.156	0.189	0.167	0.729
S03	23	85.19	0.404	0.952	2.111	1.382	0.218	0.255	0.186	0.144	0.557
S31IS3.2	13	48.15	0.298	0.972	1.519	1.371	0.196	0.222	0.186	0.115	0.615
S31IS3.4	10	37.04	0.240	0.930	1.370	1.313	0.183	0.181	0.186	-0.013	0.686
S31S5.2	19	70.37	0.449	0.926	1.778	1.588	0.328	0.320	0.186	-0.024	0.444
S31S3	12	44.44	0.286	0.957	1.444	1.377	0.194	0.216	0.186	0.098	0.625
S31S5.1	20	74.07	0.477	0.961	2.000	1.585	0.284	0.351	0.186	0.190	0.391
Z22	10	37.04	0.202	0.966	1.407	1.229	0.130	0.142	0.186	0.087	0.753
Z23	16	59.26	0.389	0.949	1.741	1.503	0.301	0.278	0.186	-0.084	0.518
Z31	10	37.04	0.253	0.957	1.407	1.344	0.174	0.187	0.186	0.072	0.674

Z33	15	55.56	0.174	0.984	1.630	1.131	0.111	0.101	0.186	-0.104	0.825
Z34	7	25.93	0.160	0.989	1.259	1.215	0.111	0.120	0.186	0.073	0.792
U22	9	33.33	0.202	0.979	1.346	1.256	0.135	0.150	0.189	0.101	0.740
U24	13	48.15	0.304	0.945	1.519	1.409	0.199	0.224	0.186	0.112	0.611
QC03	18	66.67	0.455	0.941	1.692	1.608	0.307	0.344	0.193	0.108	0.402
QC31	7	25.93	0.199	0.971	1.385	1.274	0.164	0.136	0.193	-0.203	0.764
CS441	12	44.44	0.287	0.949	1.444	1.379	0.207	0.216	0.186	0.043	0.624
CS501	6	22.22	0.148	0.992	1.231	1.194	0.139	0.111	0.189	-0.260	0.808
SC24	5	18.52	0.129	0.971	1.192	1.177	0.125	0.098	0.193	-0.274	0.830
SE802	21	77.78	0.693	0.924	2.704	1.966	0.403	0.433	0.186	0.070	0.247
SE902	16	59.26	0.326	0.943	1.852	1.356	0.194	0.208	0.186	0.067	0.638
SE111S6	11	40.74	0.226	0.965	1.423	1.276	0.115	0.164	0.193	0.297	0.715
SEG111	5	18.52	0.139	1.000	1.200	1.200	0.200	0.200	0.194	0.000	0.652
SE111S7	8	29.63	0.222	0.990	1.423	1.295	0.202	0.155	0.193	-0.302	0.730
SE412	17	62.96	0.403	0.928	1.630	1.540	0.285	0.303	0.186	0.060	0.473
SE501	9	33.33	0.216	0.974	1.333	1.285	0.102	0.162	0.186	0.372	0.718
13	12	44.44	0.320	1.000	1.462	1.462	0.462	0.308	0.189	-0.500	0.465
11	8	29.63	0.213	1.000	1.308	1.308	0.308	0.205	0.189	-0.500	0.644
17	6	22.22	0.160	1.000	1.231	1.231	0.231	0.154	0.189	-0.501	0.733
49	11	40.74	0.293	1.000	1.423	1.423	0.423	0.282	0.189	-0.500	0.510
20	5	18.52	0.124	1.000	1.185	1.170	0.167	0.117	0.186	-0.421	0.796
38	6	22.22	0.149	1.000	1.222	1.207	0.204	0.142	0.186	-0.435	0.753
86	11	40.74	0.282	1.000	1.407	1.407	0.407	0.272	0.186	-0.500	0.528

Each inbred line, if regarded as a single centroid determined according to the mean genetic similarity estimates, can be discriminated from the others on the basis of genotyping data, as shown in **Figure 2**. A clear sub-grouping of inbred lines as single centroids was obtained in the four main quadrants when the individuals belonging to each accession were plotted bidimentionally according to the principal coordinates.

A neighbor-joining (NJ) tree was also constructed on the basis of the genetic dissimilarity matrix whose mean coefficients were computed between all possible pair-wise combinations of inbred lines of the core collection (**Figure 3**).

Two main subgroups of branches with most of the accessions were generated each including about half of the inbred lines. Moreover, one of these sub-groups was further split into additional sub-nodes and two well-defined clusters wherein several inbred lines could be ordered. A few inbred lines were positioned apart from the main tree (**Figure 3**).

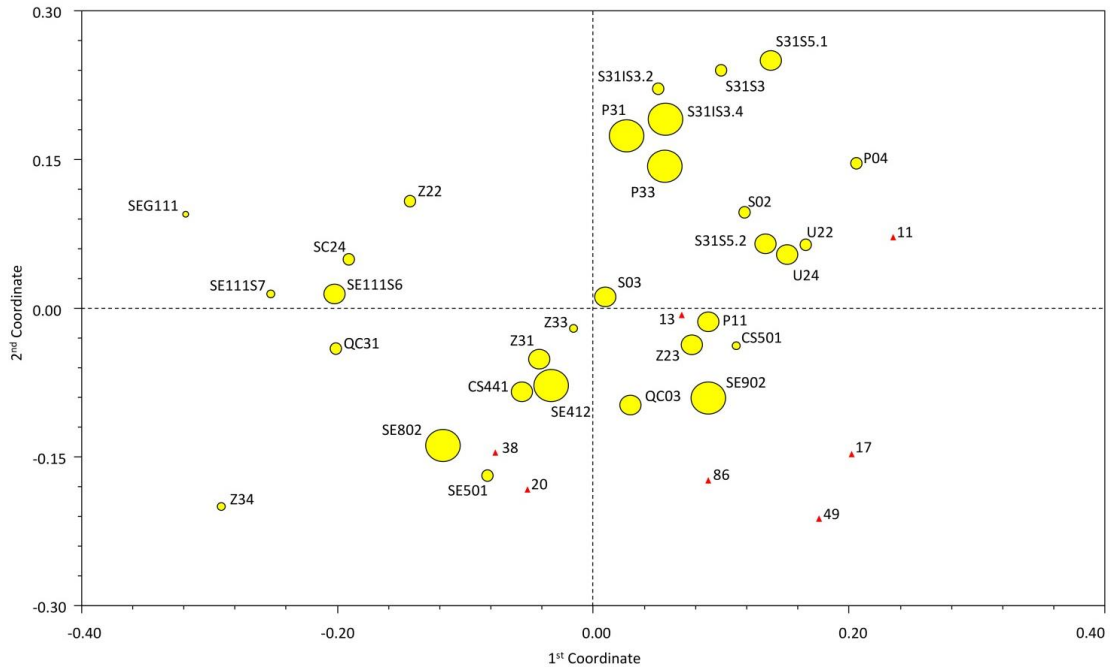


Figure 2.The centroids of all the inbred lines expressed as MGS (Mean Genetic Similarity) estimates plotted according to the first two main components. The red triangles refer to the seed parents, whereas the yellow dots indicate the pollen donors. The difference in size is related to the genetic variability found within each accession represented by the standard deviation of the Simple Matching (SM) coefficient.

The population structure was investigated using the ΔK method [11], 197 which enabled to discover two levels of genetic grouping for the inbred lines (**Appendix 1.2**). When the number of accession units (K) was set to 3 (ΔK=25), as many as 22 (59%) of the inbred lines assayed in this study were grouped in a single main cluster, while two additional small clusters were formed each represented by 6 (16%) inbred lines. Only 4 (11%) genotypes displayed an admixed ancestry (*i.e.*, membership <70%), as expected with the occurrence of genetic recombination and hybridization.

A second level of genetic structure was investigated within groups of inbred lines (**Appendix 1.2**) by setting the number of accession units (K) to 24 (ΔK=19). In this case, the whole collection of inbred lines was fragmented in clusters composed by one or few accessions (**Figure 4**). In particular, the inbred lines S02, S03 and P04 shared high membership values to a single cluster, as well as the inbred lines SE111S6, SE111S7, SEG1111, and Z33, S31S3, S31S3.2, S31S3.4 were divided in two well-defined clusters. Four additional clusters were formed by couple of inbred lines sharing high membership to

the same group. Overall, the proportion of individual genotypes that displayed admixed ancestry with membership to multiple clusters was equal to 8%.

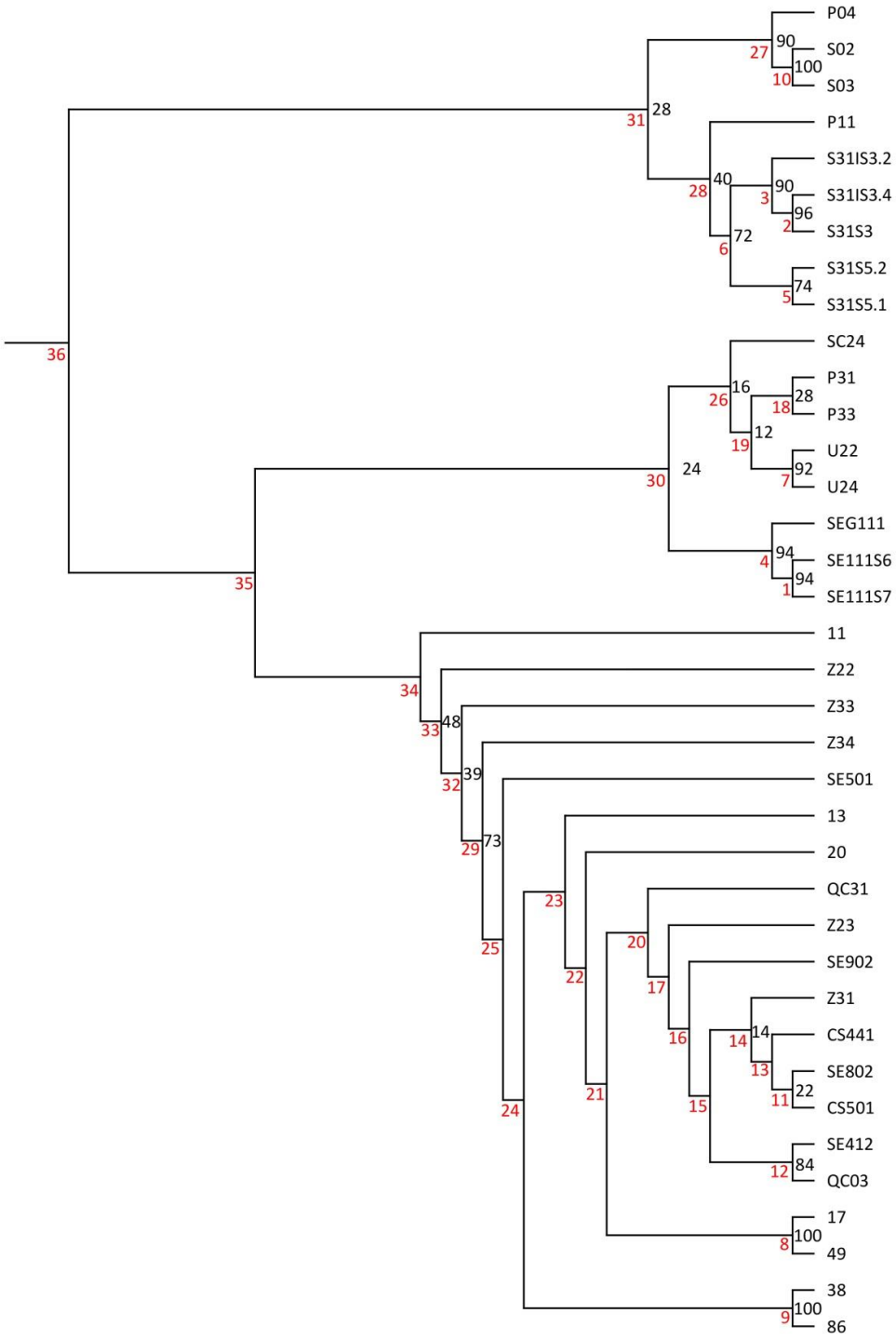


Figure 3. The NJ (Neighbor Joining) tree of the 37 inbred lines analyzed. The tree was computed using genetic dissimilarity matrix of all pair-wise comparisons between inbred lines. The numbers next to the main nodes indicate the bootstrap values (only estimates ≥30% are reported).

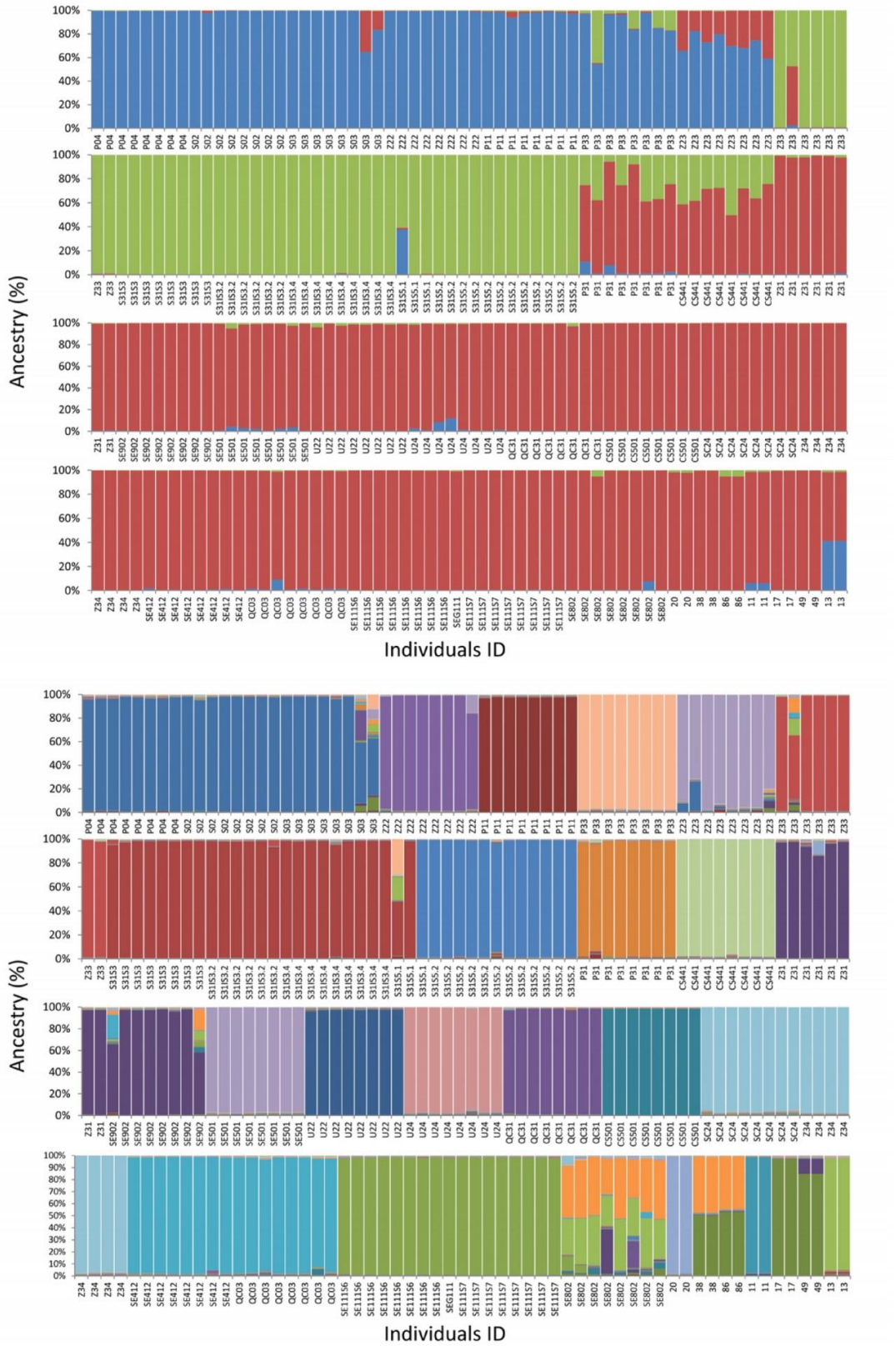


Figure 4. Estimated genetic clustering (K=3, upper panel, and K=24, lower panel) obtained with STRUCTURE. The population of inbred lines is reported on the X axis, whereas the percentage of ancestry is shown on the Y axis.

Discussion

In this study, we developed a method for genotyping elite breeding stocks of leaf chicory (*Cichorium intybus* L.) by assaying microsatellite marker loci selected for the linkage map position and polymorphism information content. The genetic identity and stability of 37 inbred lines and the extent of their genetic diversity, expressed as genetic distance and allelic divergence, were addressed by using a panel of neutral SSR markers. Furthermore, our subset of 27 mapped SSR marker loci (*i.e.*, 3 for each of the 9 basic linkage groups) proved to be very informative for the study of the population structure and inbreeding level of *C. intybus* plant materials. Along with estimates of observed homozygosity, the genetic stability of each breeding stock was addressed by computing Shannon's estimates of phenotypic diversity and Rohlf's coefficients of genetic similarity. On the whole, these statistics indicated strong genetic uniformity for the investigated inbred lines, as expected for breeding stocks developed by selfing and full-sibling programs. Furthermore, a marked inbreeding for the majority of the accessions was supported by the high homozygosity observed in nearly all the inbred lines (the mean value was as high as 78.7%). Also, the Wright's fixation index indicated that the genetic differentiation between the inbred lines is high (approximately 63%) and that one third of the genetic variation (approximately 33%) is occurring within the inbred lines, due not only to homozygosity for different marker alleles but also for heterozygosity at some of the marker loci. Our data demonstrate that most of the genetic differentiation is occurring among inbred lines; thus, each breeding stock can be considered as genetically uniform and distinguished from the others of the core collection.

It is worth mentioning that individual inbreeding coefficients, as estimate of the strength of inbreeding for single inbred lines, were shown to be low or negative, indicating that the observed heterozygosity was greater than expected. Maintenance of such levels of heterozygosity in spite of inbreeding reproductive strategies (*i.e.*, selfing, full-sibling and back-crossing) could be a consequence of the reproductive system of *C. intybus*, which is naturally characterized by high frequency of allogamy as a result of self-incompatibility. We may also speculate that a fraction of the observed heterozygosity could be a consequence of phenotypic selection (*i.e.*, morphologically superior individuals) operated by breeders during inbreeding programs.

All mapped SSR markers exploited in this multi-locus DNA genotyping method scored high polymorphism information content, with the exception of marker M8.22 that revealed an almost monomorphic condition. Although its very low or null discriminant ability, this marker locus was taken into account as it showed an allele-specific genotype, which is typical of leaf chicory and allows to identify Radicchio from other *C. intybus* types (*e.g.*, Witloof).

Regarding the NJ clustering results, the inbred lines known to be genetically related (*i.e.*, inbred lines that originated from the same local variety) proved to form a very well-defined subgroups of the tree (*e.g.*, accessions SE111 and S31). The STRUCTURE analysis of the population of genotypes (for K=25) revealed clusters of single individuals in agreement with the grouping of inbred lines shown by the NJ tree analysis. In fact, inbred lines P04, S02 and S03 were grouped in the same cluster of ancestry and associated to the same node of the tree. This is also true for several pollen donor lines, such as SEG111, SE111S6 and SE111S7, and Z34 and SC24. In addition, the inbred lines belonging to the minor clade of the NJ tree were all grouped in different STRUCTURE clusters (see **Figure 3** and **4** for details). As far as the seed parent lines (13, 20, 11, 38, 86, 17 and 49), they were grouped into four different clusters as expected on the basis of their variety of origin. Interestingly, the inbred line 13, which resulted to have an admixed ancestry, is known to originate from a specific introgression and backcross program.

PCA allowed for the definition of centroids for all the inbred lines. The first component was positively associated with cycle length, discriminating long cycle accessions from short cycle accessions, with only a very few exceptions per class. It is also worth mentioning that the centroids of inbred lines with a common origin could be plotted in different areas of the quadrants. This finding suggests that it is possible to develop and select genotypically different inbred lines (*i.e.*, homozygous of different alleles at the same loci) starting from individuals selected within a given local variety of leaf chicory, namely Radicchio of Chioggia.

Conclusions

Our research deals with the implementation and validation of a multi-locus genotyping system in leaf chicory that may have utility for the marker-assisted breeding of new varieties of Radicchio. In particular, the plant materials used in this study cover a core collection of Radicchio of Chioggia (*i.e.*, "Red of Chioggia" biotype) experimental materials, which not only manifest valuable traits but also possess applicable uses in modern breeding programs.

From a technical point of view, labeling each set of primers with different fluorescent dyes allowed us to differentiate and score up to eight SSR marker loci in a single Genescan® run. One important advantage of this method is a substantial cost savings for fluorescent primer labeling, because the synthesis of a specific fluorescently labeled primer for each SSR marker locus is not needed. The multiplex-ready PCR required only four fluorescent dye labeled primers to complete the research analyses. Furthermore, multiplex-ready PCR combines both the advantages of the M13-tailed primer method [12] and multiplex PCR [13] for fluorescent-based SSR genotyping of single individuals. The use of the M13 primer has several advantages over other techniques. First, it allows for working with a unique tail sequence and avoiding the need of using the requirement of several different SSR dye labeled primers. In addition, the technique has the further advantage of being less time consuming and reducing consumable costs. For multiplex purposes, it is only necessary to change the fluorescent colors to label the different PCR products of each SSR marker locus. We were therefore able to reconstruct the genotype of each individual across all the accessions for as many as 27 target loci (3 selected marker loci for each of the 9 linkage groups) by performing 14 PCR reactions and 4 Genescan® runs.

In conclusion, we successfully developed and implemented an efficient and reproducible method for the multi-locus genotyping of elite breeding stocks of leaf chicory belonging to the Radicchio of Chioggia biotype using 27 microsatellite marker loci scattered throughout the genome. We demonstrated that this method is useful for assessing the homozygosity and genetic stability of single inbred lines and for measuring the specific combining ability between maternal and paternal inbred lines on the basis of their genetic diversity. This information could be exploited for planning crosses and predicting the heterosis of experimental F1 hybrids on the basis of the allelic divergence and genetic distance of the

parental lines. Knowing the parental genotypes would enable not only to protect newly registered varieties but also to assess the genetic purity and identity of the seed stocks of commercial F1 hybrids, and to certificate the origin of their food derivatives.

References

[1]. Lucchin M, Varotto S, Barcaccia G, Parrini P (2008) Chicory and Endive. In Handbook of Plant Breeding, Vegetables I: Asteraceae, Brassicaceae, Chenopodicaceae. Springer Science. New York. p. 1-46. DOI:10.1007/978-0-387-30443-4_1

[2]. Barcaccia G, Tiozzo CS (2014) New male sterile mutant of leaf chicory, including Radicchio, used to produce chicory plants and seeds with traits such as male sterility exhibiting cytological phenotype with shapeless, small and shrunken microspores in dehiscent anthers. USA PATENT No. US20140157448-A1. DOI: 10.13140/2.1.3176.6400.

[3]. FAOSTAT database collections. Food and Agriculture Organization of the United Nations, Rome. http://faostat3.fao.org/home/E. Accessed 06 Apr 2015.

[4]. Barcaccia G, Pallottini L, Soattin M, Lazzarin R, Parrini P, Lucchin M (2003) Genomic DNA fingerprints as a tool for identifying cultivated types of radicchio (*Cichorium intybus* L.) from Veneto, Italy. Plant Breed. 122: 178-183. DOI:10.1046/j.1439-0523.2003.00786.x.

[5]. Zhang L, Yuan M, Tao A, Xu J, Lin L, Fang P, Qi J (2015) Genetic structure and relationship analysis of an association population in Jute (*Corchorus* spp.) evaluated by SSR markers. PLoS ONE. 10(6). DOI:10.1371.

[6]. Yunlei Z, Hongmei W, Wei C, Yunhai L, Haiyan G, Xiaohui S, Feichao H, Fanchang Z (2015) Genetic diversity and population structure of elite cotton (*Gossypium hirsutum* L.) germplasm revealed by SSR markers. Plant Syst. Evol. 301(1): 327-336. DOI:10.1007/s00606-014-1075-z.

[7]. Gonzaga ZJ, Aslam K, Septiningsih EM, Collard BCY (2015) Evaluation of SSR and SNP markers for molecular breeding in rice. Plant Breed. Biotechnol. 3(2): 139-152. DOI:10.9787/PBB.2015.3.2.139.

[8]. De Simone M, Morgante M, Lucchin M. (1997) A first linkage map of *Cichorium intybus* L. using a one-way pseudo-testcross and PCR-derived markers. Mol. Breeding. 3: 415-425. DOI:10.1023/A:1009616801404.

[9]. Van Stallen N, Vandenbussche B, Verdoodt V, De Proft M (2003) Construction of a genetic linkage map for witloof (*Chicorium intybus* L. var. *foliosum* Hegi). Plant Breed. 122: 251-525. DOI:10.1111/j.1439-0523.2003.00869.x.

[10]. Cadalen T, Mörchen M, Blassiau C, Clabaut A, Scheer I, Hilbert J-L *et al*., Quillet MC (2010) Development of SSR markers and construction of a consensus genetic map for chicory (*Cichorium intybus* L.). Mol. Breeding. 25: 699-722. DOI:10.1007/s11032-009-9369-5.

[11]. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 4: 2611-20. DOI:10.1111/j.1365-294X.2005.02553.x.

[12]. Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. A poor man's approach to genotyping for research and high-throughput diagnostics. Nat. Biotechnol. 18: 1-2. DOI:10.1038/72708.

[13]. Hayden MJ, Nguyen TM, Waterman A, Chalmers KJ (2008) Multiplex-Ready PCR: A new method for multiplexed SSR and SNP genotyping. BMC Genomics. 9: 80. DOI:10.1186/1471-2164-9-80.

[14]. Thrile TLR, Lorenz E, Frees KL, Schwartz DA. (2001) M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. Biotechniques. 31(1): 6-8.

[15]. Yeh, FC and Boyle TJB (1997) Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belg. J. Bot. 129: 157.

[16]. Levene H (1949) On a matching problem arising in genetics. Ann. Math. Stat. 20, 91-94. DOI: 10.1214/aoms/1177730093.

[17]. Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution.19: 395-420. DOI: 10.2307/2406450.

[18]. Wright S (1978) Variability Within and Among Natural Populations. Evolution and the Genetics of Populations. Vol. 4. Chicago: Univ. of Chicago Press; 1978. ISBN: 9780226910413.

[19]. Lewontin RC (1974) The genetic basis of evolutionary change. New York: Columbia University Press.

[20]. McDonald BA and McDermott JM (1993) Population genetics of plant pathogenic fungi. Bioscience. 43; 311-319. DOI: 10.2307/1312063.

[21]. Saitou N and Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406-425.

[22]. Dice LR (1945) Measurement of the amount of ecological association between species. Ecology. 26, 297-302. DOI: 10.2307/1932409.

[23]. Rohlf FJ (1988) NTSYS-PC Numerical taxonomy and multivariate analysis system. New York: Exeter Publishing, New York.

[24]. Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics. 131, 1567-1587. DOI: 10.3410/f.1015548.197423.

Chapter III

Discovery and Analysis of Nuclear Male-sterility in Radicchio, Leaf Chicory (*Cichorium intybus* subsp. *intybus* var. *foliusum* L.)

Abstract

The present research deals with the discovery and genetic analysis of nuclear recessive male-sterile mutants of leaf chicory (*Cichorium intybus* subsp. *intybus* var. *foliosum*, 2*n*=2*x*=18). These mutants, which to the best of our knowledge are the first spontaneous male-sterile mutants ever discovered and described in Radicchio, were characterized in great details for the developmental pathway of micro-sporogenesis and gametogenesis, and the inheritance pattern of the gene underlying the male-sterility trait. A quick molecular diagnostic assay was also developed for the early marker-assisted selection of the genotype associated to male-sterile plants. Overall data clearly support a nuclear origin and a monogenic control of recessive type for the male-sterility trait in each of the leaf chicory mutants. Male-gametogenesis was documented to arrest at the stage of uninucleate microspores. In particular, cytological observations revealed that microspores degenerate before their release from the tetrads, later showing a collapse of the exine. In the mutants, the totality of microspores proved to be shrunken and much smaller than wild-type ones. In fact pollen grains were never detected in mature anthers, demonstrating a full expressivity of the trait with mutants being 100% male-sterile. Moreover, the fine mapping of the mutant locus was attempted by using molecular markers. The gene responsible for male-sterility, here named *ms1*, was mapped on linkage group 4 and found tightly linked to two microsatellite loci at a genetic distance of 5.8 cM and 12.1 cM upstream and downstream the target gene. Segregation analysis of expressed sequence tags and genomic contigs encompassing predicted genes enabled to narrow down the chromosomal block to a 12.2 cM region surrounding the *ms1* locus. Overall experimental results on the male-sterile mutants are presented and their use for the breeding of new F1 hybrid varieties in red chicory is critically discussed. The discovery of non-engineered male-sterility in this species will open new frontiers for breeding new F1 varieties of Radicchio, in particular, and of chicory, in general, provided that such trait can be successfully transferred to elite inbred lines and precociously identified by molecular diagnostic assays suitable to perform marker-assisted selection.

Keywords: Chicory, male-sterility, cytogenetics, ms1 locus, molecular markers

Introduction

Chicory (*Cichorium intybus* L.) is a diploid plant species (2*n*=18), belonging to the Asteraceae family, subfamily Cichoriodeae, tribe Lactuceae or Cichorieae [1-2]. These species are naturally allogamous, due to an efficient sporophytic self-incompatibility system [3-4-5-6]. In addition, outcrossing is promoted by a floral morpho-phenology (*i.e*. proterandry, with the anthers that mature before the pistils) unfavourable to selfing in the absence of pollen donors [7-8] and by a favourable competition of allo-pollen grains and tubes (*i.e*. pollen genetically diverse from that produced by the seed parents, usually called auto-pollen) [9-10]. Long appreciated as a medical plant by ancient Greeks and the Romans [6-11], leaf chicory varieties are nowadays amongst the most important cultivated vegetable crops, being used mainly as component for fresh salads or more rarely cooked according to local traditions and alimentary habits [6]. At present, this species are grown all over continental Europe, in South Western Asia, and on limited areas in Northern America, South Africa, and Australia.

Two main groups can be recognized within *C. intybus* subsp. *intybus* to which all the cultivated types of chicory belong: the first, which refers to the var. *foliosum*, traditionally includes all the cultivar groups whose commercial products are the leaves (*i.e*. leaf chicory), while the second regards the var. *sativum* and comprises all the types whose commercial product, either destined to industrial transformation or direct human consumption, is the root (*i.e*. root chicory) as described by Lucchin *et al.* [6]. The cultivar groups of leaf chicory include mainly Witloof chicory, Pain de sucre, Catalogne and Radicchio. In particular, Radicchio is the Italian common name that has been adopted by all the most internationally used languages to indicate a very differentiated group of chicories, with red or variegated leaves, traditionally cultivated in North Eastern Italy. All the red types of Radicchio now being cultivated seem to derive from red-leaved individuals firstly introduced in XV century. According to historical information [12], the cultivation of red chicory goes back to the first half of XVI century. For sure, the original type has to be identified with the “Red of Treviso” which has been for long the only cultivated Radicchio in the Venetian territories. Originally selected around 1930, nowadays “Red of Chioggia” is by far the most widely grown among the various types of Radicchio and the one which presents the highest within-type differentiation as far as the availability of cultivars able to

guarantee an almost complete year round production. As a matter of fact, it has shown a great adaptability to very different environmental situations all around the world, becoming the most grown type of Radicchio outside the Italian country and the most known at international level [6].

It is worth mentioning that traditionally cultivated populations of leaf chicory, in general, and Radicchio, in particular, have been developed by mass selection in order to obtain uniform populations characterized by valuable production and acceptable commercial head size and shape. Newly released varieties are mainly synthetics produced by intercrossing a number of phenotypically superior plants, selected on the basis of morpho-phenological and commercial traits. More rarely, plants are also evaluated genotypically by means of progeny tests. Synthetics have a rather large genetic base and are represented by a heterogeneous mixture of highly heterozygous genotypes sharing a common gene pool. In recent years, methods for the constitution of F1 hybrids have been developed by private breeders and seed firms. Details on the procedure for the constitution of such hybrids are not available in the current literature and it may be presumed that each company has developed its own protocol, mainly in accordance to the genetic material it has at disposal and to the possibility of applying a more or less efficient control on the F1 hybrid seed production phase.

As a matter of fact, the strong self-incompatibility system, which hinders obtaining highly homozygous parents, and the absence of a male-sterility factor within the species or in sexually compatible species, made it generally difficult to propose an efficient F1 seed production scheme and, most of all, to consider these newly commercial populations or varieties as true F1 hybrids for leaf chicory.

As it happens for most allogamous species, in leaf chicory detectable heterosis effects are present and hybridization between genotypes selected on the basis of their specific combining ability gives vigorous and uniform progenies. Consequently, the constitution of F1 hybrid populations is profitable in a practical breeding scheme and it is also feasible on a large commercial scale by the selection of self-compatible genotypes, for the production of inbred lines, and the identification of genotypes showing male-sterility, to be used as seed parents for the hybridization with pollen donors. It is therefore expected that F1 hybrid populations will be bred and adopted with increasing frequency for leaf chicory, including

Radicchio. This is particularly true for the cultivated types that take a great advantage from the uniformity of the marketed products, as this is often the key for the customer's appreciation.

Male-sterility is defined as the failure of plants to produce functional anthers or pollen grains. It is more prevalent than female-sterility, likely because the male sporogenesis and gametophytes (*i.e.* pollen tubes) are less protected from environment than the female sporogenesis and gametophytes (*i.e.* embryo sacs). Male-sterile plants have propagation potentials in nature because they can still set seeds, being female-fertility unaffected by most of the mutations responsible for male-sterility. Male-sterility is known to occur spontaneously via mutations in nuclear and/or cytoplasmic genes involved in the development of anthers and pollen grains. In the model plant *Arabidopsis*, about 3,500 genes may be specifically expressed in anthers and many of them are required for the production of pollen grains [13-14]. Studying male-sterility mechanisms is immensely important not only for the understanding of reproductive barriers that act in flowering plants, but also because of their potential applications for the breeding of F1 hybrid varieties. If it is true that hybrid varieties in most of the crop species have been dramatically successful so far, it also seems likely that hybrid varieties will continue to expand in the future.

In the last century, male-sterile mutants have allowed the exploitation of heterosis (*i.e.* hybrid vigour, meant as improved characteristics in terms of size, resistance, growth rate, fertility and crop yield of a hybrid offspring over those of its parents) through the constitution of F1 hybrid varieties in many agricultural and horticultural crops. Two kinds of male-sterility can be observed in plants: nuclear and cytoplasmic male-sterility. The former type of genetic male-sterility is based solely on recessive mutations that affect different functions in nuclear genes (NMS), while cytoplasmic male-sterility (CMS) is maternally inherited and mainly due to mutations in the expression of mitochondrial genes. Moreover, in the homozygous recessive genotypes (*msms*) manifesting male-fertility, this trait can be eventually restored by nuclear-encoded fertility restorer (*Rf*) genes. In several species, nuclear and/or cytoplasmic male-sterility has been used to produce female parental lines and exploited for the production of hybrid seeds through controlled pollination with

male parental lines showing specific combining ability. In absence of an efficient genetic male-sterility system, anthers must be removed mechanically from flowers of seed parents.
Notwithstanding the high commercial interest, the presence of a naturally occurring CMS system has not been reported in leaf chicory whereas strategies to genetically engineering male-sterility were used in Magdeburg, Witloof and Chioggia genotypes [6].
In a first approach, transgenic male-sterile lines of leaf chicory were produced by expressing the ribonuclease gene RNase from *Bacillus amyloliquefaciens*, known as BARNASE, under the control of a tapetum-specific promoter originally isolated from tobacco [15]. Restorer lines for these male-sterile lines were obtained by expressing the gene coding for the so-called BARSTAR, the intracellular inhibitor of BARNASE under control of the same promoter [16-17]. The development of inbred lines and male-sterile lines provided a reliable pollination control and allowed a new hybrid seed production system, which has been registered as SeedLink™. This system for genetically engineering pollination in plants was invented and implemented by the private industry Plant Genetic Systems (Belgium).
Somatic hybridization by means of protoplast symmetric fusion between chicory and the CMS line of sunflower PET-1 was also attempted in order to promote the regeneration of interspecific hybrid plants. This kind of CMS in sunflower was identified in an interspecific cross between *Helianthus petiolaris* and *Helianthus annuus*, and it was associated with the expression of the mitochondrial gene ORF522, encoding a 15-kD polypeptide. The ORF522 gene was originated by a recombination event at the 3' of atp1 gene and its protein is detectable in flowers of CMS but not of restored lines [18-19]. The hybrid plants obtained after somatic symmetric fusion were cytoplasmic hybrids (*i.e.*, cybrids) and showed mtDNA rearrangements, indicating that symmetric fusion had the tendency to maintain the chicory mitochondrial genome. Three different kinds of sterility were observed: i) plant with anthers lacking dehiscence without, or with non-viable, pollen; ii) complete absence of the anthers; and iii) absence of both anthers and styles or the presence of reduced styles. One of these male-sterile plants was used for the production of F1 hybrids whose yields were equal to or higher that those of traditional varieties [20-21]. In a subsequent work, three different CMS chicory cybrids were backcrossed to Witloof chicory in order to transfer the male-sterile cytoplasm from an industrial chicory to a Witloof

genetic background. The transcript analysis revealed that the ORF522 is weakly expressed or not expressed at all in the cybrids. This finding led Dubreucq *et al*. [22] to conclude that ORF522 cannot be associated to the CMS observed in the chicory cybrids and to suggest that they presented a novel form of CMS, different from that of sunflower. Protoplast asymmetric fusion was used to produce male-sterile somatic hybrids between a Chioggia chicory WT accession and a PET-1 sunflower CMS line. At anthesis the regenerated cybrids had fewer and non-viable pollen grains but they could set seeds when free-pollination occurred [23]. Overall results collected so far using interspecific protoplast fusion experiments suggest that male-sterile cybrid plants can be actually regenerated in chicory. Nevertheless, it appears that mitochondrial genome re-arrangements lead to the creation of novel CMS chicory types instead of transferring the desired trait from CMS sunflower lines. The methods of transgenesis useful for making cytoplasmic male-sterile chicory plants comprising the ORF 522 of *Helianthus annuus* was patented by Delesalle *et al*. [24]. As a matter of fact, the development of inbred lines and male-sterile lines based on this biotechnological approach failed to provide any reliable hybrid seed production system in chicory.

A male-sterile mutant having a not well-defined genetic inheritance has been reported for root chicory (*Cichorium intybus* subsp. *intybus* var. *sativum*). This mutant is apparently characterized by functional male-sterility although not cytologically documented by Desprez *et al*. [25]. Recently, the use of high-density genetic linkage maps [26] allowed the fine mapping of molecular markers linked to the locus related to nuclear male-sterility, termed *NMS1*: in particular, the gene responsible for the male-sterility trait in root chicory was found associated to the linkage group 5 of chicory [27]. A different male-sterile mutation has been recently identified in leaf chicory (*Cichorium intybus* subsp. *intybus* var. *foliosum*) by Barcaccia and Tiozzo [28], and its cytological and genetic characterization is here reported. In the case of leaf chicory, the gene responsible for the male-sterility trait was found located on the linkage group 4 of chicory and the mutant locus, named *ms1*, was targeted by mapping with molecular markers [29].

Concerning markers, a few genetic studies using molecular markers have been carried out on *Cichorium* spp. mainly to characterize commercial varieties and experimental materials [30-31-32-33], to evaluate the genetic homogeneity and purity, respectively, of inbreds and

hybrids [30-34], and to investigate phylogenetic relationships between cultivars and cultivar groups of *C. intybus* and other species, both cultivated and wild, belonging to the same genus [33-35-36-37-38-39].

In recent years, three distinct saturated genetic linkage maps were constructed for chicories with different types of molecular markers, covering approximately 1,200 cM [26-27-40]. The 9 basic linkage groups were mainly based on neutral SSR markers, but many EST-derived SNP markers were also mapped. In particular, Cadalen *et al*. [26] developed a consensus genetic map, obtained after the integration and ordination of molecular marker data of one leaf (Witloof) and two root chicory progenies. Using this information, a method for genotyping elite breeding stocks of Radicchio, both local and modern varieties, by analysing mapped SSR markers possibly linked to EST-rich regions and scoring PIC>0.5, was recently developed using multiplex PCR assays [34].

It is worth emphasizing that molecular markers in *Cichorium* spp. have been exploited for selecting the mother plants of synthetics as well as for determining the distinctiveness, uniformity and stability, *i.e*. DUS testing, of newly bred varieties. In *Cichorium* spp., molecular markers should also find utility for assessing the genetic homogeneity and homozygosity of inbred lines produced by repeated selfing, measuring the genetic diversity among inbred lines in order to plan crosses and maximize heterosis in the experimental F1 hybrids, and evaluating the genetic purity and heterozygosity of seed stocks of commercial F1 hybrids.

As a new frontier, Galla *et al*. [41] reported original data from the bioinformatic assembly of the first genome draft of Radicchio, along with the most relevant findings that emerged from an extensive *de novo* gene prediction and *in silico* functional annotation of more than 18,000 unigenes.

The present work deals with the discovery and genetic analysis of male-sterile mutants of leaf chicory, belonging to the Radicchio biotype “Red of Chioggia”. These mutants, which to the best of our knowledge are the first spontaneous male-sterile mutants ever discovered and described in Radicchio, were characterized in great details for the developmental pathway of micro-sporogenesis and gametogenesis, and the inheritance pattern of the gene underlying the male-sterility trait. Moreover, the fine mapping of the mutant locus was attempted by using mapped SSR and EST markers. Experimental results on the male-sterile

mutants are presented and their use for the breeding of new F1 hybrid varieties is critically discussed.

Materials and Methods

Plant materials

Four distinct but genetically related male-sterile mutants of leaf chicory, belonging to the Radicchio biotype "Red of Chioggia", were discovered on the basis of morphological observations of anthers in experimental synthetic varieties bred by T&T® Produce, Sant'Anna di Chioggia, province of Venice, Italy (**Figure 1**).

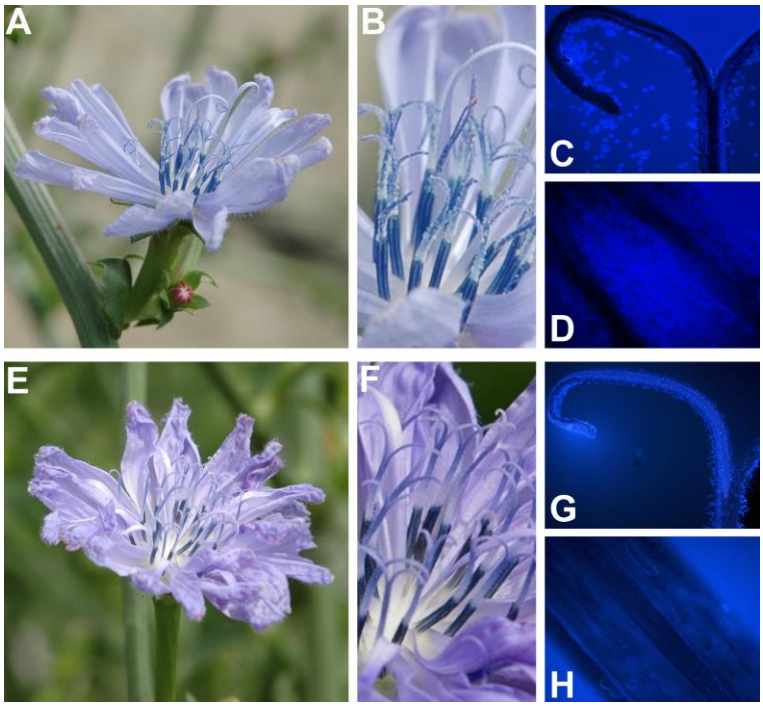


Figure 1: Phenotype of wild-type plants (A-D) and male-sterile mutants (E-H) of red chicory (*Cichorium intybus* L.). These mutants belong to experimental synthetic varieties of the biotype "Rosso di Chioggia" bred by T&T® Produce, Sant'Anna di Chioggia, province of Venice, Italy. Details of macroscopic (A, B) and microscopic (C, D) features are given for the wild-type anthers in parallel with mutant anthers (E, F and G, H, respectively).

In particular, one mutant each was found among the mother plants of synthetic lines coded as CH03/85 and CH04/99, while two mutants were found among the mother plants of CH05/01. The male-sterile mutants analyzed in this study were named L11ms, IG9ms, CS1ms and CS2ms respectively. Plant materials of *C. intybus* L. were provided by T&T® Produce (Sant'Anna di Chioggia, Venice). It is worth mentioning that the male-sterile mutants were discovered within local varieties of radicchio stemmed from recurrent

phenotypic selection programs. In particular, the three populations from which they originate have been bred through genetic selection based on progeny tests performed using mother plants chosen for uniformity and superiority of their morphological and agronomic traits. In particular, the synthetic CH05/01 is a direct derivative of CH04/99, being they bred for a different earliness (90 and 110 days, respectively), whereas the synthetic CH03/85 was bred independently for an even shorter cycle (80 days). Nevertheless, these experimental varieties are differentiated each other not only for their earliness, but also for esthetical and organoleptic traits of the commercial head.

Cytological analysis of male sporogenesis and gametogenesis in the male-sterile mutants

The presence of pollen within anthers was assayed by whole mount staining with 4',6-diamidino-2-phenylindole (Sigma Aldrich, [42]). Anther heads isolated from five flowers for each of the male-sterile mutants and the wild-types were squashed on a microscope slide and treated with 10 µl of staining solution (DAPI 5 µg/ml). After an incubation of 10 min, a detailed observation of stained anthers was done by a Leica DM4000B imagine microscope using the appropriate filter combination for DAPI fluorescent detection. Pictures were taken by the Leica DC300F camera and digital images at 10X or 20X magnification were screened in great details for the presence vs. absence of pollen grains using Adobe Photoshop® CS4 (Adobe Inc., USA, [43]).

An alternative staining technique was used to investigate the pattern of micro-sporogenesis and the development of pollen grains in each male-sterile mutant in comparison with wild-type. Flowers at four different developmental stages, spanning from young buds to full anthesis, were collected from mutants and wild-type plants, fixed in Carnoy's solution (ethyl alcohol-acetic acid 3:1) and stored at +4°C for 24-48 hours. After this pre-treatment, flowers were transferred in 70% ethyl alcohol at +4°C until their use for cytological analysis. Anthers were dissected from individual flowers, opened on microscope slides using a pair of teasing needles with the aid of a stereomicroscope. Specimens containing pollen mother cells, tetrads, microspores and pollen grains were squashed using a drop of 4% aceto-carmine and mounted in lacto-phenol with acid fuchsin.

For the preparation of meiocyte chromosomes, anther specimens of mutants and wild-types were treated with citrate buffer (10 mM citric acid, 10 mM sodium citrate, pH 4.5) for 3 min and incubated in a six times diluted pectolytic enzyme mixture containing 1% pectolyase Y23, 1% cellulase RS and 1% cytohelicase (Sigma Aldrich, [42]) in 10 mM citrate buffer at 37°C for about 1-2 hours, according to the anther stage. Anther preparations were squashed on microscope slides using a drop of purified and deionized water (Milli-Q Integral Water Purification System, [44]) and then transferred on a hot plate at 45°C. Cells were spread on microscope slides using a teasing needle by adding one drop of 45% acetic acid, then maintained at 45°C for 2 min and washed with Carnoy's solution. Each slide was dried on the hot plate at 45°C and specimens were stained with DAPI.
Cytological observations of male meiosis and gametogenesis as well as karyological analysis of meiocyte chromosomes were made under natural and fluorescent light using a photomicroscope (Zeiss Axiophot photomicroscope, [45]) equipped with epifluorescence illumination and single-band filters for DAPI. Photograph films were scanned at 1,200 dpi for digital image processing with Adobe Photoshop® CS4 (Adobe Inc., U.S.A. [43]).

Genetic analysis of mutants and inheritance pattern of the male-sterility trait

Each male-sterile mutant plant was crossed as seed parent with a wild type pollinator belonging to the same Radicchio variety. Several F1 plants from each hybrid population were then selfed and crossed in pair-wise combinations in order to obtain segregating F2 progenies. Moreover, F1 plants were also backcrossed as pollen donors with either male-sterile mutants belonging to F2 progenies or wild type plants of S1 progenies stemmed from selfing in order to obtain segregating BC1 progenies (**Figure 2**). A total of 118 F2 and 92 BC1 individuals belonging to the Red of Chioggia biotype were obtained and used for mapping the *ms* locus using AFLP and SSR markers. The F2 plants derived from the progenies of mutants CS1ms and CS2ms, whereas the BC1 progenies included plants of the mutant L11ms.

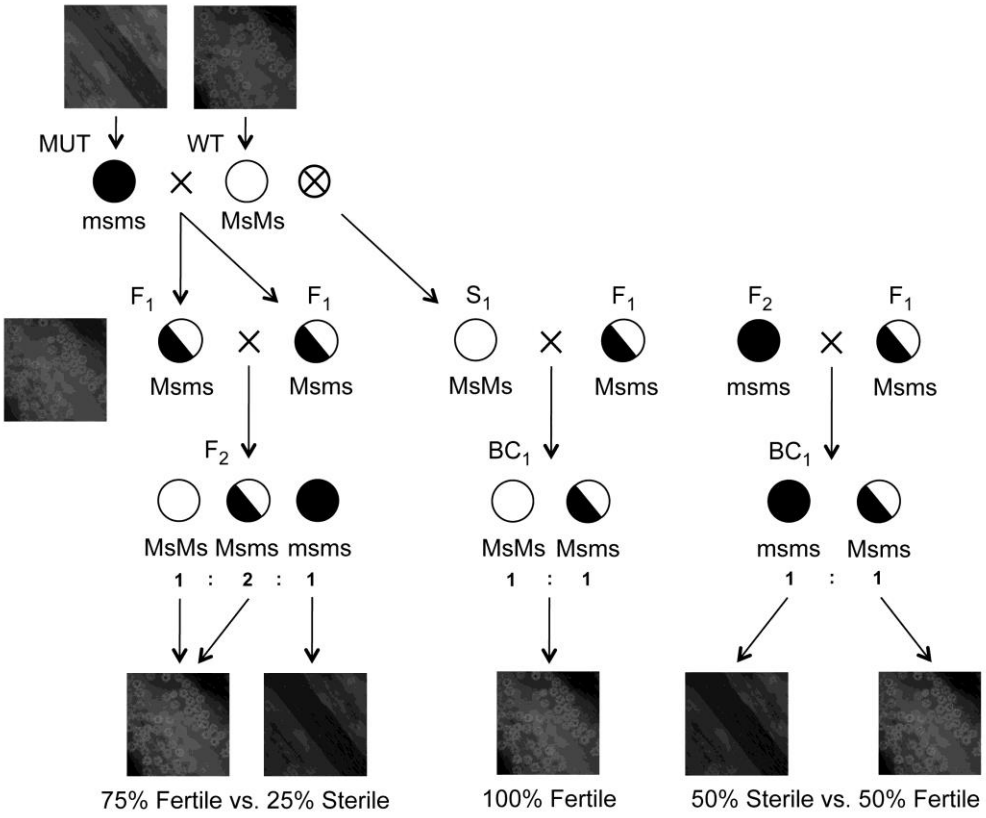


Figure 2: Genetic analysis of male-sterile mutants based on segregation patterns observed in F2 and BC1 progenies. Each of the male-sterile mutants was crossed as seed parent with a wild type pollinator belonging to the same variety. Several F1 plants from each hybrid population were then selfed and crossed in pair-wise combinations in order to obtain segregating F2 progenies. Moreover, F1 plants were also backcrossed as pollen donors with either male-sterile mutants belonging to F2 progenies or wild type plants of S1 progenies stemmed from selfing in order to obtain segregating BC1 progenies. These experimental populations were used to establish the inheritance pattern of the mutation and to map the male-sterility gene using microsatellite markers.

Moreover, 100 plants belonging to an F1 population obtained by pseudo-testcross and segregating for male-sterility were produced from the fourth mutant IG9ms. All these segregating populations were analyzed for understanding the genetic control of male-sterility and for a preliminary screening of molecular markers co-segregating with male-sterility (**Table 1**).

Table 1. List of primer pairs used to amplify the mapped markers of LG4. The name of the locus, the GenBank accession of the trait amplified, the polymorphism observed in ms mutants and the primer pairs themselves are mentioned.

Locus Name	GenBank ID	Polymorphism	Primer Pairs	References
EU02M09[3]	JF748831	$(TC)_n$	For: GGCATCGGGATAGAAAAACA Rev: TCAATGCCTCAACAGAAATCC	[3]Barcaccia and Tiozzo (2014)
EU03H01[1,3] **M4.11b**[2]	KF880802	$(TG)_nCG(TG)_n$	For: GCCATTCCTTTCAAGAGCAG Rev: AACCCAAAACCGCAACAATA	[1]Cadalen *et al.* (2010) [2]Ghedina *et al.* (2015) [3]Barcaccia and Tiozzo (2014)
EU07G10[1] **M4.10b**[2]	KX534081	$CT)_nCATA$ $(CA)_nCT(CA)_n$	For: CATCCATTATTGGGCAG Rev: CACCAACGAACTCCTTACAAA	[1]Cadalen *et al.* (2010) [2]Ghedina *et al.* (2015)
MADS box L2/R2[1] **CAPS marker**[4]	AF101420[1] KX45584/0/1[4]	*Nco*I	For: TTTTTGTGGGGTTTTGATTTTAGA Rev: TGAGATTGCATGAATGAGAACA	[1]Cadalen *et al.* (2010) [4]Present study

In addition, a BC1 population of 198 individuals was bred as follows. An F1 male-fertile plant, heterozygous at the male-sterility locus (*i.e.*, *Msms*) obtained by crossing a male-sterile plant (i.e., *msms*) belonging to a cultivated population of Radicchio (Red of Chioggia, CH04/99) with a wild accession (*i.e.*, *MsMs*) was backcrossed to the same male-sterile plant previously adopted and maintained by cuttings. This latter population was used for the fine mapping of the male-sterility gene by means of mapped SSR markers and EST clones deriving from Cadalen *et al.* [26]), Barcaccia and Tiozzo [28-29], Ghedina *et al.* [34] and Galla *et al*. [41].

Molecular mapping of the gene for male-sterility in Radicchio

Total genomic DNA was isolated from 100 mg of fresh leaf tissue using the DNeasy® Plant mini-kit (QIAGEN, [46]) following the recommendations of the manufacturer. Quality and concentration of DNA samples were estimated by spectrophotometric analysis (NanoDrop 2000c UV-Vis, Thermo Fisher Scientific) and agarose gel electrophoresis (1.0% w/v agarose TAE 1X gel containing 1X SYBR® Safe, Thermo Fisher Scientific).

A subset of 48 progeny plants with a contrasting microgametogenesis pattern, (*i.e.* 24 male-sterile plants and 24 male-sterile plants) were selected and used for performing a bulked segregant analysis [47]. Genomic DNA bulks of 12 plants each from two progeny sets were prepared by combining equal amounts of DNA from male-fertile and male-sterile plants.

All bulked DNA samples were investigated by AFLP markers using the parental lines as controls. Genomic AFLP fingerprinting was performed using the protocol of Vos *et al*. [48] with modifications described by Barcaccia *et al*. [33]. AFLP analysis was based on the detection of *Eco*RI-*Mse*I genomic restriction fragments by PCR amplification with 9 different primer combinations having three selective nucleotides (E+CAC, E+CCA, E+CTG and M+ATC, M+AGG and M+AAG), chosen during preliminary tests according to their ability to find homologous binding sites in red chicory templates. Overall data were recorded as a binary matrix by assigning the molecular weight to each quantitatively polymorphic marker identified by comparing DNA fingerprints with known DNA ladders.

The AFLP-derived amplicon corresponding to the marker E02M09/230 (see **Table 1)** was recovered from the agarose gels, subcloned into plasmid vectors and sequenced in order to obtain information on the whole genomic sequence. The SCAR (*i.e.* Sequence Characterized Amplified Region) marker developed from the AFLP amplicon E02M09/230 was then used to develop the diagnostic DNA marker E02M09/163. As a preliminary screening based on mapped SSR markers, 12 male-sterile and 12 male-fertile plants were randomly selected from segregating populations of each mutant, for a total of 96 genomic DNA samples. Microsatellite (SSR) loci analysis was carried out as described by Hayden *et al*. [49] and Ghedina *et al*. [34]. For each linkage group, 1 to 4 mapped marker loci were analyzed. The analysis of DNA fragments was carried out using a fully automated capillary electrophoresis system (Applied Biosystems 3130) and SSR patterns were visualized and scored in replicated analysis using the software GeneScan® v. 2.1 e Genotyper® v. 2.0 (Applied Biosystems). Following this preliminary test, genomic loci providing marker alleles co-segregating with the male-sterility/fertility phenotypes were investigated in the whole collection of segregating populations (n=300).

The observed segregation ratio of SSR markers was tested by chi-square analyses for goodness-of-fit to the expected 3:1 or 1:1 segregation ratios, as well as for independent assortment in the male-sterile vs. wild-type progenies by a 2×2 contingency test. Segregation data for the markers were analyzed with JoinMap® v. 2.0 [50] using the cross pollination (CP) population type option. The association between molecular markers and male-sterility locus was assessed by recording the target *ms1* locus as a putative gene fully co-segregating with the trait being mapped. For the identification of the linkage group

carrying the *ms* locus with the selected SSR markers, the grouping module was applied by setting a minimum LOD = 3 and a maximum recombination frequency, *r* = 30% [51]. The genetic distance between each pair-wise comparison of SSR marker locus and *ms* locus, expressed in centiMorgans (cM), was calculated from the recombination frequency corrected by using the Kosambi mapping function [52].

According to the genetic linkage map developed for *C. intybus* by Cadalen *et al.* [26], a candidate gene for male-sterility, corresponding to a MADS-box gene (AF101420) was mapped in the linkage group 4. Specific primer pairs designed using PerlPrimer v1.1.21 (Source Forge, [53]) were used to amplify the full-length sequence and sub-regions of the MADS-box gene. Genomic DNA sequences of the MADS-box gene were amplified from male-sterile and male-fertile plants (KX455840 and KX455841, respectively) and investigated to find SNPs possibly associated to male-sterility and male-fertility. Amplification reactions were performed in a 9700 Thermal Cycler (Applied Biosystems) with the following conditions: 94°C for 5 min followed by 30 cycles at 94°C for 30 sec, 57°C for 30 sec, 72°C for 60 sec and a final extension of 10 min at 72°C and then held at 4°C. The quality of PCR products was assessed by electrophoresis on 2% (w/v) Agarose gel stained with 1X SYBR® Safe™ DNA Gel Stain (Life Technologies) and visualized by KODAK 1D Image Analysis Software. Following amplification reactions, PCR products were restricted by using the enzyme *Nco*I (Promega, [54]) following the protocol suggested by the manufacturer. The restriction reaction was incubated at 37°C for 2 hours. Quality of the enzymatic reaction was assessed by 2.5% (w/v) Agarose gel (Life Technologies) stained with 1X SYBR® Safe™ DNA Gel Stain (Life Technologies) gel electrophoresis and visualized by KODAK 1D Image Analysis Software.

Genetic mapping with molecular marker

The marker alleles that proved to co-segregate with the male-sterility/fertility phenotypes in the preliminary screenings were further investigated using a total of 198 individuals of the BC1 mapping population obtained by crossing a cultivated accession and a wild individual (as described above).

Segregation data were analysed with JoinMap® v. 2.0 [50] using the backcross (BC1) population type option. The association between microsatellite markers and the male-

sterility locus was assessed by recording the target *ms1* locus as a putative gene fully co-segregating with the trait being mapped. For the genotype code option, the presence of one marker allele and wild-type phenotype were assigned to a homozygous (=M_1M_1) or a heterozygous (=M_1M_2) locus, whereas the presence of another marker allele and male-sterile phenotype to the other homozygous (=M_2M_2) locus. For the identification of the linkage group carrying the *ms* locus with the selected SSR markers, the grouping module was applied by setting a minimum LOD = 3 and a maximum recombination frequency, r = 40%. The genetic distance between each pair-wise comparison of marker locus and target locus, expressed in centiMorgans (cM), was calculated from the recombination frequency corrected by using the Kosambi's mapping function Kosambi [52].

Markers annotation on the genomic contigs linked to the ms locus

The map of the linkage group carrying the *ms* locus was enriched by annotating each marker (*i.e.* subjects) on the first genome draft (*i.e.* queries) of leaf chicory [41] by means a default BLASTN approach. Matches scoring E-values included from 3.00 E-36 to 0 are presented. Moreover, putative functional annotation of genes included in these contigs was performed with a BLASTX approach on TAIR10 database. The sequence of all mapped markers were deposited in GenBank.

Results

Cytogenetic analysis of male sterile mutants

Male-sterility of the four distinct but genetically related leaf chicory mutants presented in this study is controlled by a nuclear gene that acts as recessive. All crosses between male-sterile mutants and wild-type pollinators resulted in 100% male-fertile F1 progenies, whereas F2 and BC1 progenies showed to segregate for this trait and to be composed of both male-fertile and male-sterile plants, with proportions equal to 3:1 and 1:1, respectively. Segregation ratios observed in the F2 and BC1 progenies developed for each of the four male-sterile mutants along with chi-square values are reported in **Table 2**.

Table 2. Segregation ratios observed in the F2 and BC1 populations bred for each of the male-sterile mutants along with chi-square values.

Mutants	Progeny type	Progeny size	Expected ratios		Observed ratios		Chi-square values
			male-fertile plants	male-sterile plants	male-fertile plants	male-sterile plants	
CS1ms	F2	107	80	27	82	25	0.153
CS2ms	F2	92	69	23	71	21	0.232
IG9ms	F2	100	75	25	78	22	0.480
L11ms	F2	84	63	21	66	18	0.571
Overall	F2	383	287	96	297	86	1.324
CS1ms	BC1	94	47	47	49	45	0.170
CS2ms	BC1	102	51	51	54	48	0.353
IG9ms	BC1	88	44	44	41	47	0.409
L11ms	BC1	96	48	48	43	53	1.042
Overall	BC1	380	190	190	187	193	0.095

Overall data clearly supported a nuclear origin and a monogenic control of recessive type for the male-sterility trait in each of the red chicory mutants. Taking together all segregating progeny sets of the F2 and BC1 populations, which included 383 and 380 plants respectively, chi-squares values were non-significant, being as low as 1.324 and 0.095 (**Table 2**).

For what concerns male sporogenesis and gametogenesis, regular meiosis was normally found in wild-type plants. After meiosis, each microspore of the tetrads was shown to develop into a binucleate pollen grain through a mitotic division that originated a vegetative and a generative nucleus. Moreover, at anthesis when the pollen grains were mature, they

germinated and emitted the pollen tubes (*i.e.* the microgametophyte), in which the generative nucleus underwent another mitotic division, giving rise to two distinct sperm nuclei.

In the male-sterile mutants, the cytological analysis performed with aceto-carmine showed that microsporogenesis proceeded normally up to the development of microspore tetrads. Then the microspores arrested their development at the uninucleate stage, as documented in **Figure 3** using a parallel with unrelated wild-type plants. In particular, cytological observations revealed that microspores degenerate before their release from the tetrads showing a collapse of the exine. At the end of male meiosis, most of the microspores were found arranged in tetrads while some others were released, becoming shapeless even though the cytoplasm stained well with aceto-carmine. At the beginning of gametogenesis, non-viable shrunken microspores were clearly visible within anthers (details given on **Figure 3**).

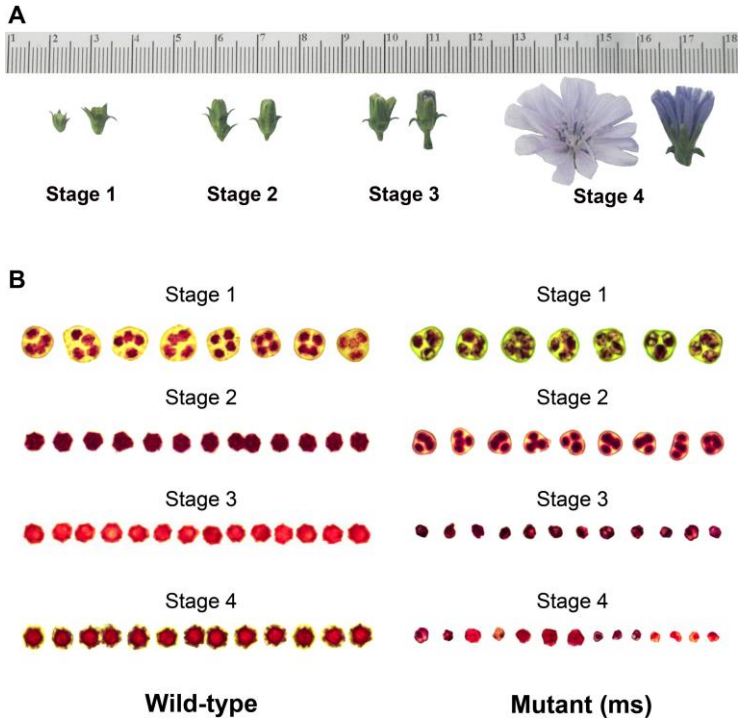


Figure 3: Flower developmental stages in red chicory (panel A) and patterns of male gametogenesis in the male-sterile mutants in parallel with wild-type plants (panel B). In the mutants, the microspores of each tetrad arrest their development at the uninucleate stage, degenerating before their release from the tetrads. At full flowering, most of the microspores of dehiscent anthers were found shapeless, shrunken and much smaller than wild-type ones.

It is worth mentioning that microspore tetrads were comparable for their size and shape between mutants and wild-types, whereas mutant microspores at the uninucleate stage proved to be shrunken and much smaller than wild-type ones, as shown on **Figure 4**.

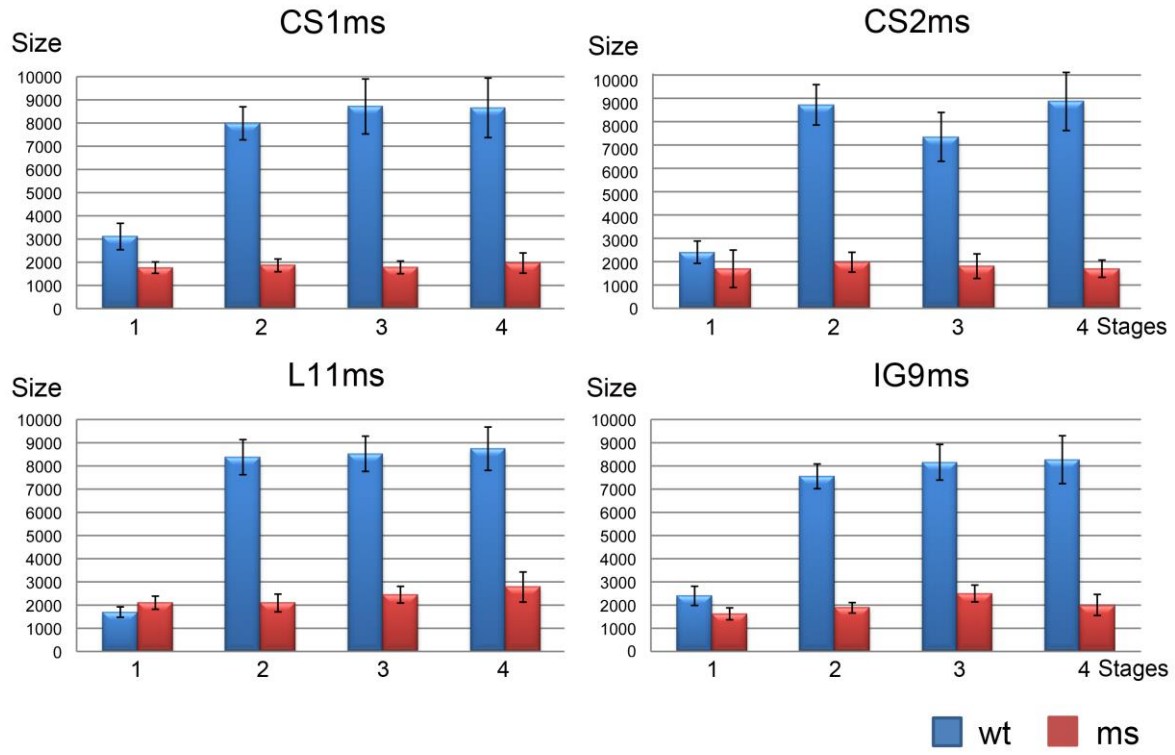


Figure 4: Microspore size in the male-sterile mutants compared to wild-type plants, expressed as mean value (histograms) with standard error (bars). At the stage of tetrad, the microspores were comparable for their size and shape between mutants and wild-types, whereas mutant microspores at the uninucleate stage proved to be about three times smaller than wild-type ones.

The most important evidence is that pollen grains were never detected in mature anthers of all four male sterile mutants. This cytological finding was also supported by DAPI staining of squashed anthers (see **Figure 1**, panels C-D and G-H).

Furthermore, the cytological analysis of microsporogenesis and gametogenesis was performed in the plants belonging to F2 and BC1 progenies. At the cellular level, male meiosis was shown to proceed regularly until the stage of microspore tetrads in both male-sterile mutants and male-fertile plants. Gametogenesis followed a regular pathway in male-fertile plants, giving rise to mature pollen grains, whereas microspores collapsed within each tetrad in the male-sterile plants, without any further developing process (**Figure 5**). In

fact, at the end of gametogenesis, a similar phenotype of non-viable shrunken microspores was observed for male-sterile mutants belonging to each of the segregating progenies (**Figure 5**).

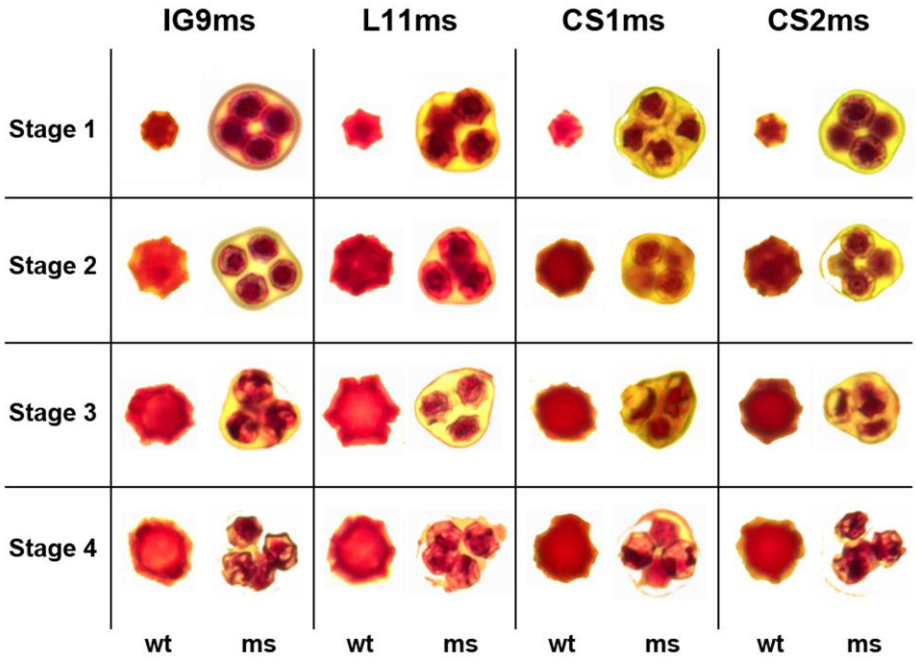


Figure 5: Parallel between male gametogenesis in wild-type plants and male-sterile mutants belonging to segregating progenies at flower stages 1-4. Gametogenesis followed a regular pathway in male-fertile plants, giving rise to mature pollen grains, whereas microspores collapsed within each tetrad in the male-sterile plants, without any further developing process. This finding demonstrated that the gene responsible for male-sterility is inherited in the offspring from each mutant by recovering an unaltered maternal genotype, which is always associated to an unchanged phenotype for male-sterility.

The chromosome behavior of male-sterile mutants was also investigated during meiosis: the male meiocyte chromosomes were further analyzed by means of DAPI staining in both wild-type and mutant flowers. Different forms of meiotic abnormalities were found in the male-sterile mutants compared to wild-types, especially at prophase I. In fact, during pachytene, the stage when chiasmata take place and crossing-over occurs between non-sister chromatids of homologous chromosomes, abnormal pairings and chromosomal loops were observed in several sites. Moreover, chromatin bridges were also observed in ana-telophase II. **Figure 6** shows some examples of normal chromosome pairing in wild-types

(panel A) and miss-pairing of certain chromosome pairs in male-sterile mutants (panels B-D).

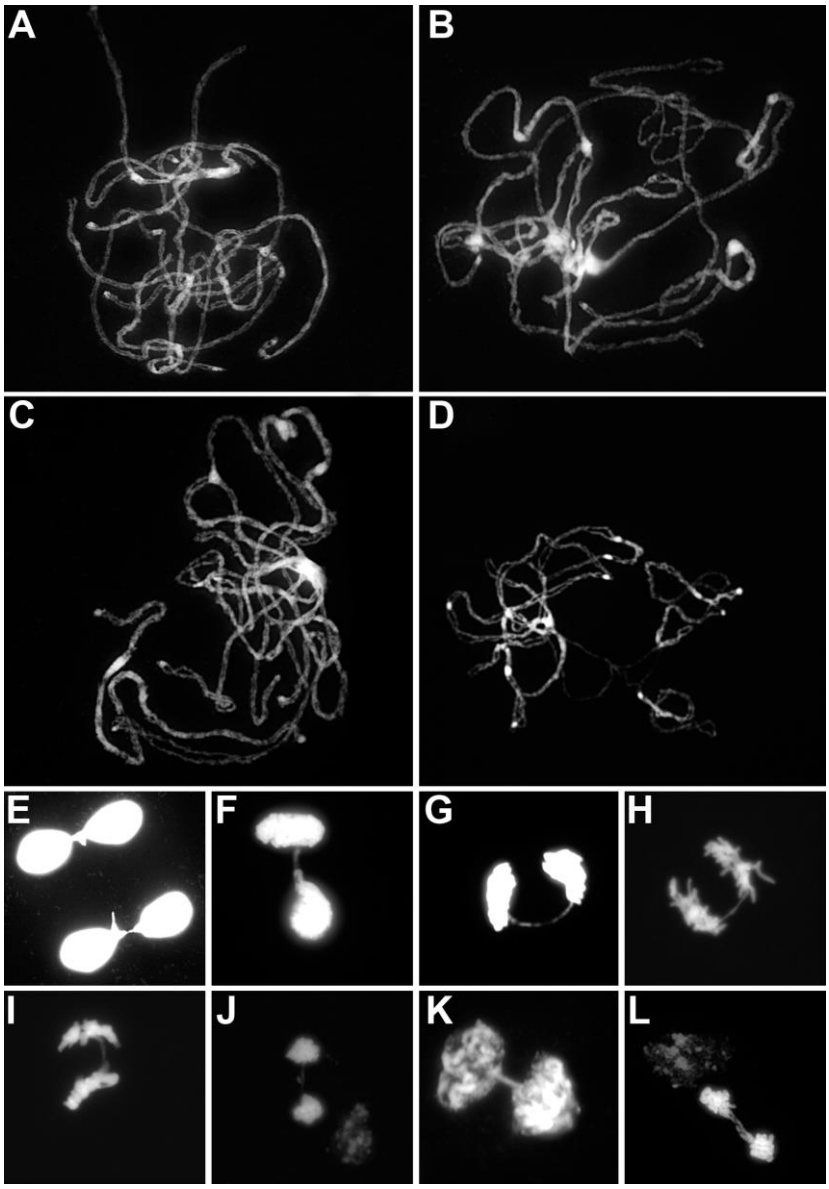


Figure 6: Results of cytogenetic analyses of male-sterile mutants: different types of meiotic abnormalities were found in the male-sterile mutants compared to wild-types, especially at prophase I, along with chromatin bridges observed in ana-telophase II. Some examples of normal chromosome pairing in wild-types (panel A) and miss-pairing of certain chromosome pairs in male-sterile mutants (panels B-D). The main aberrant feature in the mutants was recovered at pachytene stage when the homologous chromosomes reached their full pairing: homologues were not completely pairing each other and aberrant structures characterized by one or more loops, due to partial or aspecific pairing between homologous chromosomes, were often observed (see panels B-D). Moreover, several cases of chromatin bridges, *i.e*. bridges made of chromatin occurring between newly forming cells, were found in the male-sterile mutants (panels E-L).

The main aberrant feature was recovered at pachytene stage when the homologous chromosomes reached their full pairing. It was evident that in the mutant, the homologues were not completely pairing each other and aberrant structures characterized by one or more loops, due to partial or aspecific pairing between homologous chromosomes, were often observed (see white arrows in panels B-D of **Figure 6**). Moreover, several cases of chromatin bridges, *i.e*. bridges made of chromatin occurring between newly forming cells, were found in the male-sterile mutants (**Figure 6**, panels E-L).

Molecular mapping of the male sterility locus

In order to map the *ms1* locus, a subset of F2 progenies was initially screened to find out molecular marker alleles co-segregating with the male-sterility/fertility trait. This strategy allowed us to obtain informative AFLP fingerprints and to detect few DNA markers qualitatively polymorphic between DNA bulks of male fertile and male sterile progeny plants. One of the AFLP markers of interest proved to encompass a microsatellite showing a perfect dinucleotide repetition of the motif (TC/GA)n, with n ranging from 27 to 33. As a consequence, this information was used to convert it into a single-locus SCAR marker, specifically an SSR-based assay, for the detection of its marker alleles. This microsatellite region was shown to include the basic dinucleotide repeat TC/GA, ranging in size from 141 to 171 bp in relation to the genotypes analysed. The SSR marker locus was coded as E02M09/230 (GenBank accession JF748831). In the BC1 progenies, the SSR genotypes represented by the short marker alleles at the homozygous state was associated to the male-sterile *msms* individuals, whereas the homozygous state for the long marker alleles and the heterozygous condition were related to the male-fertile *MsMs* and *Msms* individuals, respectively.

When the datasets for both the trait and the marker were analyzed together, there was a significant deviation in the segregation data from the expected 1:1:1:1 ratio. The genetic determinant for male-sterility was found tightly associated with the diagnostic marker, as their alleles were preferentially inherited together (Fisher's 2×2 contingency test: $\chi 2$=75.3 with P<0.0001). However, recombination events were apparently possible in the chromosome block carrying the male-sterility gene. In fact, this gene was associated with the marker E02M09 [28], renamed M4.12 by Ghedina *et al*. [29], in a chromosome window

likely characterized by active crossing-over sites and densely saturated by expressed sequence tags.

The assignment of the *ms1* gene to the linkage group 4 of the consensus genetic map of *Cichorium intybus* [26] was obtained by testing the co-segregation of mapped molecular markers with the mutant phenotype. Among the microsatellite markers publicly available for the chicory genome [26], the SSR locus coded as EU03H01, renamed M4.11b by Ghedina *et al*. [34], containing an imperfect microsatellite motif (TG)nCG(TG)n, with total n varying up to 11 (KF880802), was found associated to the linkage group 4 of the consensus map and thus to the locus for male-sterility [29]. The fine genetic mapping of these selected markers (**Table 1**) was pursued and the genetic recombination estimates were validated using 198 BC1 individuals of the segregating population on the basis of chi-square values against independent assortment patterns. The genetic distance between the male-sterility locus and the microsatellite markers mapped upstream and downstream the *ms1* gene was equal to 5.8 cM and 12.1 cM, respectively (**Figure 7**).

An additional mapped microsatellite marker belonging to linkage group 4, corresponding to the SSR coded as EU07G10 [26] renamed as M4.10b [34] was also analysed. This SSR marker (GenBank accession KX534081) mapped at 31.2 cM away from the *ms1* locus, the trait responsible of male sterility.

Concerning the MADS-box L2/R21 gene (AF101420), the alignment of sequences corresponding to its 5'-UTR region, exon 1 and the early region of intron 1 were recovered from both mutants and wild types plants and enabled to map the MADS-box locus on the linkage group 4 by means a CAPS markers. In fact, three SNPs were discovered in the amplified sequence of the first exon of the MADS-box gene. Remarkably, the restriction site of the six-base cutter *Nco*I enzyme included a polymorphism at position 61 of the nucleotide sequence of the male fertile genotype when compared with the male sterile genotype (GenBank accessions KX455840 and KX455841), thus allowing the analysis of such polymorphism as a CAPS marker. The amplification-restriction protocol for the CAPS marker was applied to the total 198 BC1 individuals of the mapping population. In particular, 14 individuals out of the total analyzed showed recombinant genotypes for the male-sterile/fertile phenotypes with an assessed genetic distance of 7.3 cM (**Figure 7**).

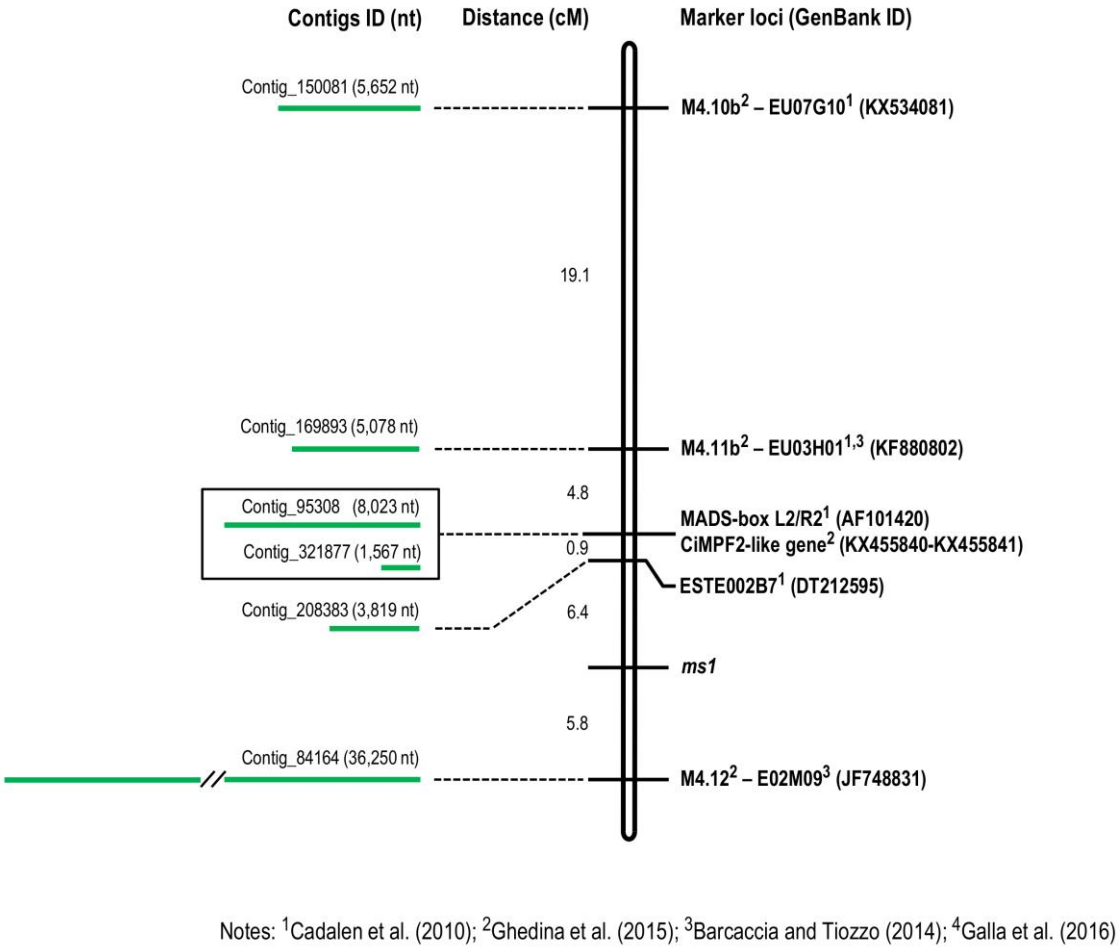


Figure 7: Schematic representation of part of the linkage group 4 (*i.e.* LG4) of *Cichorium intybus* consensus map containing DNA markers described in **Table 1** and the *ms1* locus where the mutant gene responsible for male sterility is located in Radicchio. Genomic contigs of the first genome draft of leaf chicory obtained by Galla *et al.* [41] that matches with the mapped markers are reported together with map distances in cM and GenBank accessions of the markers.

Data published by Galla *et al.* [41] on the genome draft of leaf chicory were used to discover, predict and annotate all genes of the genomic contigs encompassing the molecular markers mapped on the linkage group 4. All the mapped markers could match with at least one genomic contig of the genome draft (for details see **Figure 7** and **Table 3**).

Concerning the marker E02M09, its sequence was found to match with Contig_84164, long 36,250 nucleotides and putatively carrying genes sharing high similarity with the Arabidopsis genes AT3G11330, AT5G62230, AT5G50170 and AT5G63130 (**Table 3**). It

is worth noting that the former gene , namely AT3G11330, encodes for a PIRL-like protein, a member of the Plant Intracellular Ras-group-related LRRs (*i.e.*, Leucine-rich repeat proteins) and results to be required for differentiation of microspores into pollen. Furthermore, the longest genomic contig matching with the CAPS marker developed in this study was Contig_95308, long 8,023 nucleotides and containing the gene model AT4G24540 which encodes for a MPF2-like protein involved in flowering (**Table 3**).

On the basis of the consensus genetic map presented by Cadalen *et al.* [26], the ESTE002B7 was placed in its most likely position 0.9 cM downstream the MADS-box (**Figure 7**).

Table 3. List of markers, related genomic contigs of the first genome draft [41] and hypothetic function on the bases of matches with *Arabidopsis* protein database (TAIR10).

Marker Map Reference	Query GenBank ID	Marker GenBank ID	E-value	Gene Model Name	Predicted Function
E02M09[3]	84164_7	JF748831[3]	4,00E-95	AT5G50170	C2 calcium/lipid-binding and GRAM domain containing protein
	84164_8		1,00E-76	AT3G11330	Plant intracellular ras group-related LLR (PIRL)
	84164_1		3,00E-29	AT5G63130	Octicosapeptide/Phox/Bem1p family-like protein
	84164_6		9,00E-08	AT5G62230	ERECTA like protein
MADs box L2/R2[1]	95308	AF101420[1]	4,00E-26	AT4G24540	Agamous-like 24 (AGL24)
CAPS marker[4]	321877	KX5455840/1[4]	2,00E-25	AT2G22540	SVP like protein
EU03H01[1,3] **M4.11b**[2]	169893	KF880802[1,2,3]	1,00E-51	AT1G01620	Plasma membrane intrinsic like protein
EU07G10[1] **M4.10b**[2]	150081	KX534081[4]	1,00E-58	AT5G39000	Malectin/receptor-like protein kinase family protein
ESTE002B7[1]	208383	DT212595[1]	1,00E-75	AT1G58360	Amino acid permease like protein

[1]Cadalen *et al.* [26]
[2]Ghedina *et al.* [34]
[3]Barcaccia and Tiozzo [29]
[4]Present study

Discussion and Conclusions

Male sterility has been reported in more than 610 plant species [55]; it includes cytoplasmic male sterility (CMS), which is caused by mitochondrial genes with coupled nuclear genes, and nuclear male sterility (NMS), which is caused by nuclear genes alone. The male-sterile phenotype can manifest itself as varying reproductive abnormalities [57-58-59-60]. In some cases, male reproductive organs (*e.g.* stamens) are transformed into petals or female reproductive organs (*e.g.* carpels) [61-62]. Other male-sterile mutations lead to the degeneration of anthers or developing pollen grains that fail to develop fully, and if they do develop completely, they are often not functional [63]. In some instances, *e.g.* sunflower, petunia and maize, anthers are often completely missing [64]. Because male sterile phenotypes encompass a large variety of reproductive abnormalities [57-58], different genes may be involved in, or responsible for male-sterility.

Dozens of CMS and NMS systems have been studied at the genetic and molecular levels [65]. A detailed list of CMS and respective restorer (*Rf*) genes was reviewed by Chen and Liu [65] in major crops as corn, rice, sunflower, brassicas, radish, sorghum, wheat, common bean, pepper, carrot and sugar beet. In contrast to CMS, most NMS mutants are less suitable for hybrid seed production because their male-sterility traits cannot be efficiently maintained [65]. Consequently, it is not surprising that genes responsible of CMS are much more studied than ones causing NMS.

A deep comprehension of male-sterility systems is extremely important in plant breeding applications, being male-sterility largely used as one of the most effective methods to produce F1 hybrids in crop plants.

F1 hybrids are usually constituted by crossing two highly homozygous parental lines selected in order to obtain so a highly heterozygous progeny, which is usually characterized not only by high uniformity of phenotypic traits, but also by strong heterosis in terms of productivity. The best way to breed F1 hybrid varieties is based on a female line, seed parent, unable to produce viable pollen hence avoiding self-fertilization. Commercial F1 hybrid breeding programs require an efficient tool to control pollination during seed production in a large-scale [66]. Hence, the use of male-sterility to produce F1 hybrid seed has proven to be cost-effective and it has been widely exploited in some major crops such as maize [67], sorghum [55], rice [68], rapeseed [69], rye [70], wheat and pearl millet [71-

72]. In maize, the most prominent example of F1 hybrid breeding crop, many of the commercially used varieties are produced by CMS [67].

Molecular marker technology allows to map gene of agronomic importance in segregating generations by comparison of near isogenic lines [73] or by bulked segregant analysis (BSA) [47]. Using different molecular marker techniques, several NMS genes in plants have been mapped. Examples include the rice (*Oryza sativa* L.) photoperiod-sensitive NMS genes [74-75-76-77], rice thermo-sensitive NMS gene *tms5* [78], the Chinese cabbage (*Brassica campestris* L. ssp. *chinensis*) NMS gene *gms* [79], the soybean (*Glycine max* L. Merr.) NMS gene *ms* [80], the tomato (*Lycopersicon esculentum* Mill.) NMS gene *ms14* [81], and the sunflower (*Helianthus. annuus* L.) NMS genes *ms10* and *ms11* [82].

Among Compositae, the family which leaf chicory belongs, sunflower (*Helianthus annuus* L.) is the most important crop in the word. In sunflower, NMS is mostly controlled by single recessive genes [83]. NMS P21 (*ms11*) and NMS HA 89–552 (*ms7*) were adopted for sunflower breeding program [84-85]. Of all the sunflower NMS genes mapped, *ms10* of B11A3 was mapped to LG11 with flanking RFLP markers at distances of 4.8 and 9.8 cM; *ms11* of NMS P21 was mapped to LG8 with flanking SSR markers at genetic distances of 3.8 and 4.1 cM and *ms9* of NMS HA 89–360 was mapped to LG10 with flanking markers at distances of 1.2 cM on both sides [86]. Gong *et al*. [87] mapped the *ms6* gene of NMS HA 89–872 to LG16 with flanking markers at distances of 7.2 and 18.5 cM, *ms7* to LG6 with flanking makers at distances of 2.6 and 4.7 cM, and *ms8* to LG5 with flanking markers at distances of 7.4 and 3.8 cM.

To the best of our knowledge, the presence of a naturally occurring CMS system has not been reported in chicory [6], whereas just one spontaneous NMS mutation has been discovered and exploited by breeders in root chicory [88-89].

Recently, the gene *NMS1* proved to cause male sterility in root chicory, with mutant individuals showing visibly lack of pollen and shorter anthers if compared with the wild-type ones. It is worth noting that, the gene NMS1 was mapped to a 0.8 cM region on LG5 [27].

Consequently, this is the first time that naturally occurring male-sterile mutants are discovered and genetically characterized in leaf chicory. In all male-sterile mutants, the cytological analysis performed with aceto-carmine showed that microsporogenesis proceed

regularly up to the development of tetrads, the microspores then arrest their developmental program. At the beginning of microgametogenesis, non-viable shrunken microspores were clearly visible within anthers. Conversely, detailed cytogenetic investigations indicated the occurrence of meiotic abnormalities in the male-sterile mutants, especially at prophase I. In fact, abnormal pairings and chromosomal loops were observed during pachytene. It is well known that the central function of synapsis is the recognition of homologues by pairing, an essential step for a successful meiosis. Irregular synapsis for some of the homologous chromosomes may alter the further development of microspores, leading to the failure of gametogenesis.

Although nuclear male-sterile mutations can affect both microsporogenesis and microgametogenesis and both mutation types hinder the formation of pollen grains, the genetic factors affecting meiotic chromosome pairing in plants are of special interest to geneticists and especially breeders. Several lines of evidence indicate that a crucial step for male fertility is the conjugation of chromosomes during the first meiotic prophase, a phenomenon termed synapsis. When synapsis occurs irregularly in the male meiocytes, meiosis may lead to non-functional microspores. Several mutants characterized by the lack of chromosome pairing during the first meiotic prophase (*i.e.*, asynapsis) have been found in plant species, as well as mutants in which chromosomes initially pair in early meiotic prophase but fail to remain paired at later meiotic stages (*i.e.*, desynapsis) [89-90].

Our findings suggest that the exchange of DNA segments over regions of homology is strongly prevented in the male-sterile mutants and that the lack of regular synapsis for some of the homologous chromosomes may alter the further development of microspores. In addition, the occurrence of chromatin bridges between newly forming cells is usually an indicator of abnormalities related to cellular division. All together these features provide karyological evidences that support chromosome features and factors negatively influencing the process of male gametogenesis, resulting in the absence of properly formed pollen grains observed in our mutants.

It is worth noting that the sterility of gametes occurs only in male organs. The quantity of seeds set by the mutant flowers was not significantly different from that of wild-type plants, demonstrating that the female organs of mutant flowers are completely fertile. As a consequence, our observations suggest that the mutant phenotype is attributable to a gene

expressed in meiocytes and required for normal progression of meiosis and/or release of microspores from the tetrad.

The gene responsible for male-sterility was found genetically linked to the genomic locus M4.12 (JF748831), an AFLP marker encompassing as SSR region, about 5.8 cM apart from the *ms1* locus. Two additional SSR markers corresponding to genomic loci M4.10b (KX534081) and M4.11b (KF880802) were found genetically associated to the *ms1* locus on linkage group 4, according to the genetic map developed by Cadalen *et al.* [26]. Moreover, a diagnostic CAPS marker derived from MADS-box L2/R2 gene (KX455840-KX455841) was found linked at 7.3 cM from the *ms1* locus.

According to Cadalen *et al.* [26], the SSR markers mapped on linkage group 4 by Barcaccia and Tiozzo [28-29] encompass a genomic block that includes a MADS-box similar to an MPF2 gene. Functional analyses by molecular genetic studies in model eudicots, such as *A. thaliana* L., have shown that transcription factors encoded by these genes are essential for the regulation of various aspects of flower development [91-92]. In particular, gene function studies based on silenced lines of *Arabidopsis* and spontaneous mutants of rice suggested that *CiMPF2*-like is a good candidate for male-sterility. Despite of our expectations, mutations that we came across on the MADS-box gene could not be responsible for male-sterility in leaf chicory. The polymorphism found on this expressed sequence was however useful to develop a CAPS marker associated to the *ms1* locus on linkage group 4. The sequence of this latter marker was uncovered on a rather long genomic contig of the genome draft of leaf chicory (*i.e.*, Contig_84164). Bioinformatics allowed us to reveal that this contig contains a gene whose transcripts putatively encode for a PIRL-like protein, a member of the Plant Intracellular Ras-group-related LRRs (*i.e.*, Leucine-rich repeat proteins), that is required for differentiation of microspores into pollen.

In conclusion, the male sterility gene was mapped on a well saturated and characterized linkage group in a chromosomal region spanning 7.3 cM and 5.8 cM from the *ms1* locus. On the whole, this information was crucial to plan a Genotyping-by-Sequencing experiment based on BC1 progenies with the aim of narrowing down the genomic window containing the gene for male sterility in leaf chicory.

References

[1]. Bremer K (1994) Asteraceae: cladistics and classification. Timber Press, Portland, USA.

[2]. Panero JL and Funk VA (2002) Toward a phylogenetic subfamilial classification for the Compositae. Proc. Biol. Soc. Wash., 115: 909-922.

[3]. Eenink AH (1981) Compatibility and incompatibility in witloof - Chicory (*Cichorium intybus* L.). 2. The incompatibility system. Euphytica. 30: 77-85. DOI:10.1007/BF00033662.

[4]. Varotto S, Pizzoli L, Lucchin M, Parrini P (1995) The incompatibility system in Italian red Chicory (*Cichorium intybus* L.). Plant Breed. 114: 535-538. DOI:10.1111/j.1439-0523.1995.tb00851.x.

[5]. Barcaccia G, Varotto S, Soattin M, Lucchin M, Parrini P (2003) Genetic and molecular studies of sporophytic self-incompatibility in *Cichorium intybus* L. Proc. of the Eucarpia meeting on Leafy Vegetables Genetics and Breeding, March 19-21, 2003, Noordwijkerhout, The Netherland, p. 154.

[6]. Lucchin M, Varotto S, Barcaccia G, Parrini P (2008) Chicory and Endive. In Handbook of Plant Breeding, Vegetables I: Asteraceae, Brassicaceae, Chenopodicaceae. Springer Science. New York. p. 1-46. DOI:10.1007/978-0-387-30443-4_1.

[7]. Pécaut P (1962) Etude sur le système de reproduction de l'endive (*Cichorium intybus* L.). Ann. Amélior. Plantes. 12: 265-296.

[8]. Desprez BF, Delesalle L, Dhellemmes C, Desprez MF (1994) Genetics and breeding of industrial chicory. Comptes Rendus de l'Académie d'Agriculture de France. 80(7): 47-62. ISSN: 0989-6988.

[9]. Desprez FF and Bannerot H (1980) A study of pollen tube growth in witloof chicory. In: Maxon Smith JW, Langton A (eds.) Proc. of the Eucarpia meeting on leafy vegetables, Littlehampton. pp. 47-52.

[10]. Eenink AH (1982) Compatibility and incompatibility in witloof-chicory (*Cichorium intybus* L.). 3. Gametic competition after mixed pollinations and double pollinations. Euphytica, 31: 773-786.

[11]. Wittop Koning DA and Leroux A (1972) La chicoré e dans l'histoire de la médecine et dans la céramique pharmaceutique. Rev. Hist. Pharm. (Suppl.) 215: pp. 28.

[12]. Bianchedi A (1961) I radicchi di Treviso: storia, coltivazione, forzatura e commercio. Ital. Agric. 98: 37-51.

[13]. Goldberg RB, Beals T, Sanders PM (1993) Anther development: basic principles and practical applications. Plant Cell. 5: 1217-1229. DOI:10.1105/tpc.5.10.1217.

[14]. Ma H (2005) Molecular genetic analyses of microsporogenesis and microgametogenesis in flowering plants. Annu. Rev. Plant Biol. 56: 393-434. DOI:10.1146/annurev.arplant.55.031903.141717.

[15]. Mariani C, De Beuckeleer M, Trueltner J, Leemans J, Goldberg RB (1990) Induction of male sterility in plants by a chimaeric ribonuclease gene. Nature. 347: 737-741. DOI:10.1038/347737a0.

[16]. Denis M, Delourne R, Gourret JP, Mariani C, Renerd M (1993) Expression of engineered nuclear male sterility in *Brassica napus*: genetics, morphology and sensitivity to temperature. Plant Phys. 101(4): 1295-1304.

[17]. Reynaerts A, Van de Wiele H, de Sutter G, Janssens J (1993) Engineered genes for fertility control and their application in hybrid seed production. Sci. Hort. 55: 125-139. DOI:10.1016/0304-4238(93)90028-O.

[18]. Horn R, Köhler RH, Zetsche K (1991) A mitochondrial 16-kDA protein is associated with cytoplasmic male sterility in sunflower. Plant Mol. Biol. 17: 29-36. DOI:10.1007/BF00036803.

[19]. Monegèr F and Smart CJ (1994) Nuclear restoration of cytoplasmic male sterility in sunflower is associated with the tissue-specific regulation of a novel mitochondrial gene. EMBO J. 13(1): 8-17.

[20]. Rambaud C, Dubois J, Vasseur J (1993) Male-sterile chicory cybrids obtained by intergeneric protoplast fusion. Theor. Appl. Gen. 87(3): 347-352. DOI:10.1007/BF01184922.

[21]. Rambaud C, Bellamy A, Dubreucq A, Bourquin JC, Vasseur J (1997) Molecular analysis of the fourth progeny of plants derived from a cytoplasmic male sterile chicory cybrid. Plant Breed. 116(5): 481-486. DOI:10.1111/j.1439-0523.1997.tb01035.x.

[22]. Dubreucq A, Berthe B, Asset JF, Boulidard L, Budar F, Vasseur J, Rambaud C (1999) Analyses of mitochondrial DNA structure and expression in three cytoplasmic male-sterile chicories originating from somatic hybridisation between fertile chicory and CMS sunflower protoplasts. Theor. Appl. Gen. 99(7-8): 1094-1105. DOI:10.1007/s001220051314.

[23]. Varotto S, Nenz E, Lucchin M, Parrini P (2001) Production of asymmetric somatic hybrid plants between *Cichorium intybus* and *Helianthus annuus.* Theor. Appl. Gen. 102: 6/7, pp. 950-956. DOI:10.1007/s001220000485.

[24]. Delesalle L, Dhellemmes C, Desprez M, inventors; Florimond Desprez Veuve Et Fils, assignee (2004) Methods of making cytoplasmic male sterile chicory plants comprising the ORF 522 of *Helianthus annuus*. United States patent US 6803497.

[25]. Desprez BF, Delesalle L, Dhellemmes C and Desprez MF (1994) Génétique et amélioration de la chicorée industrielle. CR Acad. Agr. Fr. 80(7): 47-62). Patented in the

République Française on 1 February 2002 by NUNHEMS ZADEN BV (Stérilité male de legumes de Cichorium cultivé et utilisation pour la production de semences hybrides, No de publication: FR2832290.

[26]. Cadalen T, Mörchen M, Blassiau C, Clabaut A, Scheer I, Hilbert J-L *et al*., Quillet MC (2010) Development of SSR markers and construction of a consensus genetic map for chicory (*Cichorium intybus* L.). Mol. Breeding. 25: 699-722. DOI:10.1007/s11032-009-9369-5.

[27]. Gonthier L, Blassiau C, Mörchen M, Cadalen T, Poiret M, Hendriks T, Quillet MC (2013) High-density genetic maps for loci involved in nuclear male sterility (*NMS1*) and sporophytic self-incompatibility (*S*-locus) in chicory (*Cichorium intybus* L., Asteraceae). Theor. Appl. Genet. 126(8), 2103-2121. DOI:10.1007/s00122-013-2122-9.

[28]. Barcaccia G and Tiozzo CS (2012) New male sterile Cichorium spp. mutant, parts or derivatives, where male sterility is due to a recessive nuclear mutation linked to a polymorphic molecular marker, useful for producing F1 hybrids of Cichorium spp. EU PATENT No. WO2012163389-A1. DOI: 1013140/2.1.4749.5044.

[29]. Barcaccia G and Tiozzo CS (2014) New male sterile mutant of leaf chicory, including Radicchio, used to produce chicory plants and seeds with traits such as male sterility exhibiting cytological phenotype with shapeless, small and shrunken microspores in dehiscent anthers. USA PATENT No. US20140157448-A1. DOI: 10.13140/2.1.3176.6400.

[30]. Bellamy A, Vedel F, Bannerot H (1996) Varietal identification in *Cichorium intybus* L. and determination of genetic purity of F1 hybrid seed samples, based on RAPD markers. Plant Breed. 115: 128-32. DOI:10.1111/j.1439-0523.1996.tb00887.x.

[31]. Koch G and Jung C (1997) Phylogenetic relationships of industrial Chicory varieties revealed by RAPDs and AFLPs. Agronomie. 17: 323-333. https://hal.archives-ouvertes.fr/hal-00885847/document

[32]. Van Stallen N, Noten V, Demeulemeester M, De Proft M (2000) Identification of commercial Chicory cultivars for hydroponic forcing and their phylogenetic relationships revealed by random amplified polymorphic DNAs and amplified fragment length polymorphisms. Plant Breed. 119: 265-270. DOI:10.1046/j.1439-0523.2000.00499.x.

[33]. Barcaccia G, Pallottini L, Soattin M, Lazzarin R, Parrini P, Lucchin M (2003) Genomic DNA fingerprints as a tool for identifying cultivated types of radicchio (*Cichorium intybus* L.) from Veneto, Italy. Plant Breed. 122: 178-183. DOI:10.1046/j.1439-0523.2003.00786.x.

[34]. Ghedina A, Galla G, Cadalen T, Hilbert JL, Tiozzo CS, Barcaccia G (2015) A method for genotyping elite breeding stocks of leaf chicory (*Cichorium intybus* L.) by assaying mapped microsatellite marker loci. BMC Res. Notes. 8(1) : 1-12. DOI:10.1186/s13104-015-1819-z.

[35]. Vermeulen A, Desprez B, Lancelin D, Bannerot H (1994) Relationships among *Cichorium* species and related genera as determined by analysis of mitochondrial RFLPs. Theor. Appl. Genet. 88: 159-166. DOI:10.1007/BF00225892.

[36]. Kiers AM, Mes TH, Van Der Meijden R, Bachmann K (2000) A search for diagnostic AFLP markers in *Cichorium* species with emphasis on Endive and Chicory cultivar groups. Genome. 43(3): 470-6. DOI:10.1139/g00-024

[37]. Van Stallen N, Noten V, Neefs V, De Proft M (2001) The phylogenetic relationship between different *Cichorium intybus* cultivars and cultivar groups, as revealed by RAPDs. Plant Breed. 120: 425-428. DOI:10.1046/j.1439-0523.2001.00633.x.

[38]. Hodgins KA, Lai Z, Oliveira LO, Still DW, Scascitelli M, Barker MS et al. and Burke JM (2014) Genomics of Compositae crops: reference transcriptome assemblies and evidence of hybridization with wild relatives. Mol. Ecol. Resour. 14(1): 166-177. DOI:0.1111/1755-0998.12163.

[39]. Raulier P, Maudoux O, Notté C, Draye X, Bertin P (2015) Exploration of genetic diversity within *Cichorium endivia* and *Cichorium intybus* with focus on the gene pool of industrial chicory. Genet. Resour. Crop. Ev. 1-17. DOI 10.1007/s10722-015-0244-4.

[40]. Muys C, Thienpont CN, Dauchot N, Maudoux O, Draye X, Van Cutsem P (2013) Integration of AFLPs, SSRs and SNPs markers into a new genetic map of industrial chicory (*Cichorium intybus* L. var. *sativum*). Plant Breed. 113(1): 130-137. DOI:10.1111/pbr.12113.

[41]. Galla G, Ghedina A, Tiozzo CS, Barcaccia G (2016) Toward a first high-quality genome draft for marker-assisted breeding in leaf chicory, Radicchio (*Cichorium intybus* L.). In Plant Genomics, Dr. Ibrokhim Y. Abdurakhmonov (Ed.), InTech. DOI:10.5772/61747.

[42]. Sigma Aldrich. Available from: http://www.sigmaaldrich.com

[43]. Adobe Photoshop® CS4. Available from http://www.adobe.com/it/products/photoshop

[44]. Milli-Q Integral Water Purification System. Avalable from http://www.millipore.com

[45]. Zeiss Axiophot photomicroscope. Available from http://www.zeiss.com

[46]. DNeasy® Plant mini-kit, QIAGEN. Available from http://www.qiagen.com

[47]. Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. USA 88: 9828–9832. DOI:10.1073/pnas.88.21.9828.

[48]. Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: A new technique for DNA fingerprinting. Nucleic Acids Res. 23: 4407–4414. DOI:10.1093/nar/23.21.4407.

[49]. Hayden MJ, Nguyen TM, Whatman A, McMichael GL, Chalmers KJ (2008) Application of multiplex-ready PCR for fluorescence-based SSR genotyping in barley and wheat. Mol. Breeding. 21: 271-281. DOI:10.1186/1471-2164-9-80.

[50]. Stam P and Van Ooijen JW (1995) JOINMAP™ version 2.0: Software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen, The Netherlands.

[51]. Barcaccia G, Albertini E, Rosellini D, Tavoletti S, Veronesi F (2000) Inheritance and mapping of 2n egg production in diploid alfalfa. Genome, 43: 528–537.

[52]. Kosambi DD (1944). The estimation of map distances from recombination values. Annals Eugenics. 12: 172-175. DOI:10.1111/j.1469-1809.1943.tb02321.x.

[53]. PerlPrimer v1.1.21. Available from: http://perlprimer.sourceforge.net

[54]. Promega. Available from http://www.promega.com

[55]. Kaul MLH (1988) Male Sterility in Higher Plants. SpringerVerlag: Berlin. DOl:10.1007/978-3-642-83139-3.

[56]. Vedel F, Pla M, Vitart V, Gutierres S, Ch´etrit P, De Paepe R (1994) Molecular basis of nuclear and cytoplasmic male sterility in higher plants. Plant Physiol. Biochem. 32: 601-8.

[57]. Laser KD and Lersten NR (1972) Anatomy and cytology of microsporogenesis in cytoplasmic male sterile angiosperms. Bot. Rev. 38, 425—454. DOI:10.1007/BF02860010.

[58]. Schnable PS and Wise RP (1998) The molecular basis of cytoplasmic male sterility and fertility restoration. Trends Plant Sci. 3: 175–80. DOI:10.1016/S1360-1385(98)01235-7

[59]. Hanson MR and Bentolila S (2004) Interactions of mitochondrial and nuclear genes that affect male gametophyte development. Plant Cell 16:S154–69. DOI:10.1105/tpc.015966.

[60]. Carlsson J, Leino M, Sohlberg J, Sundstroem JF, Glimelius K (2008) Mitochondrial regulation of flower development. Mitochondrion 8: 74—86. DOI:10.1016/j.mito.2007.09.006.

[61]. Zubko MK (2004) Mitochondrial tuning fork in nuclear homeotic functions. Trends Plant Sci. 9: 61—64. DOI:10.1016/j.tplants.2003.12.001.

[62]. Linke B and Borner T (2005) Mitochondrial effects on flower and pollen development. Mitochondrion 5: 389—402. DOI:10.1016/j.mito.2005.10.001.

[63]. Chase CD (2007) Cytoplasmic male sterility: a window to the world of plant mitochondrial–nuclear interactions. Trends Genet. 23: 81–90. DOI:10.1016/j.tig.2006.12.004.

[64]. Rieseberg LH and Blackman BK (2010) Speciation genes in plants. Ann. Bot. 106, 439—455. DOI:10.1093/aob/mcq126

[65]. Chen L and Liu YG (2014) Male sterility and fertility restoration in crops. Annu. Rev. Plant Biol. 65: 579-606. DOI:10.1146/annurev-arplant-050213-040119

[66] Horn R, Friedt W (1999) CMS sources in sunflower: different origin but same mechanism?. Theor. Appl. Genet. 1;98(2): 195-201. DOI:10.1007/s001220051058.

[67]. Acquaah G (2012) Breeding hybrid cultivars. In: Principles of Plant Genetics and Breeding (2nd Edition), 355—373. John Wiley and Sons Ltd, Chichester, UK. DOI:10.1002/9781118313718.ch18.

[68]. Barclay A (2010) Hybridizing the world. Rice Today 9, 32—35. International Rice Research Institute (IRRI), Philippines.

[69]. Zhao TJ and Gai JY (2006) Discovery of new male-sterile cytoplasm sources and development of a new cytoplasmic-nuclear malesterile line NJCMS3A in soybean. Euphytica 152: 387—396. DOI:10.1007/s10681-006-9226-0

[70]. Geiger HH and Schnell FW (1970) Cytoplasmic male sterility in rye (*Secale cereale* L). Crop Sci. 10: 590—593.

[71]. Rajeshwari R, Sivaramakrishnan S, Smith RL, Subrahmanyam NL (1994) RFLP analysis of mitochondrial DNA from cytoplasmic male-sterile lines of pearl millet. Theor. Appl. Genet. 88: 441 - 448. DOI:10.1007/BF00223658

[72]. Havey MJ (2004) The use of cytoplasmic male sterility for hybrid seed production. In: H. Daniell, and C. D. Chase (eds), Molecular Biology and Biotechnology of Plant Organelles, 623—634. Springer, Berlin Heidelberg, New York. DOI:10.1007/978-1-4020-3166-3_23.

[73]. Muehlbauer GJ, Specht JE, Thomas-Comton MA, Staswick PE, Bernard RL (1988) Near-isogenic lines—a potential source in the integration of conventional and molecular marker linkage maps. Crop Sci 28: 729–735. DOI:10.2135/cropsci1988.0011183X002800050002x.

[74]. Subudhi PK, Borkakati RP, Virmani SS, Huang N (1997) Molecular mapping of a thermo-sensitive genetic male sterility gene in rice using bulked segregant analysis. Genome 40: 188–194. DOI:10.1139/G09-092.

[75]. Wang B, Xu WW, Wang JZ, Wu W, Zheng HG, Yang ZY, Ray JD, Nguyen HT (1995) Tagging and mapping the thermo-sensitive genic male-sterile gene in rice (*Oryza sativa* L.) with molecular markers. Theor. Appl. Genet. 91: 1111–1114. DOI:10.1007/BF00223928

[76]. Zhang Q, Shen BZ, Dai XK, Mei MH, Saghai Maroof MA, Li ZB (1994) Using bulked extremes and recessive class to map genes for photoperiod sensitive genic male sterility in rice. Proc Natl Acad Sci USA 91:8675–8679.

[77]. Koh HJ, Son YH, Heu MH, Lee HS, McCouch SR (1999) Molecular mapping of a new genic male-sterility gene causing chalky endosperm in rice (*Oryza sativa* L.). Euphytica 106:57–62. DOI:10.1023/A:1003575016035.

[78]. Wang YG, Xing QH, Deng QY, Liang FS , Yuan LP, Weng ML, Wang B (2003) Fine mapping of the rice thermo-sensitive genic male-sterile gene tms5. Theor. Appl. Genet. 107: 917–921. DOI:10.1007/s00122-003-1327-8.

[79]. Miao Y, Dreyer F, Cai A, Jung C (2003) Molecular markers for genic male sterility in Chinese cabbage. Euphytica 132:227–234. DOI:10.1023/A:1024642717416.

[80]. Jin W, Palmer RG, Horner HT, Shoemaker RC (1998) Molecular mapping of a male-sterile gene in soybean. Crop Sci 38: 1681–1685. DOI:10.2135/cropsci1998.0011183X003800060043x.

[81]. Gorman SW, Banasiak D, Fairley C, McCormick S (1996) A 610 kb YAC clone harbors 7 cM of tomato (*Lycopersicon esculentum*) DNA that includes the male sterile 14 gene and a hotspot for recombination. Mol. Gen. Genet. 251: 52–59. DOI:10.1007/BF02174344.

[82]. Perez-Vich B, Berry ST, Velasco L, Fernández-Martínez JM, Gandhi S, Freeman C, Heesacker A, Knapp SJ, Leon AJ (2005) Molecular mapping of nuclear male sterility genes in sunflower. Crop science. 1;45(5): 1851-7. DOI:10.2135/cropsci2004.0684

[83]. Jan CC (1992) Inheritance and allelism of mitomycin C- and streptomycin-induced recessive genes for male sterility in cultivated sunflower. Crop. Sci. 32: 317–320. DOI:10.2135/cropsci1992.0011183X003200020006x.

[84]. Beard HB (1982) Registration of *Helianthus* germplasm pools I; II; III, IV. Crop Sci 22:1276-1277

[85]. Jan CC (1992) Cytoplasmic–nuclear gene interaction for plant vigor in *Helianthus* species. Crop Sci. 32: 320–323. DOI:10.2135/cropsci1992.0011183X003200020007x.

[86]. Chen J, Hu J, Vick BA, Jan CC (2006) Molecular mapping of a nuclear male-sterility gene in sunflower (*Helianthus annuus* L.) using TRAP and SSR markers. Theor. Appl. Gen. 113(1): 122-127. DOI:10.1007/s00122-006-0278-2.

[87]. Gong L, Li C, Capatana A, Feng J, Qi L, Seiler GJ, Jan CC (2014) Molecular mapping of three nuclear male sterility mutant genes in cultivated sunflower (*Helianthus annuus* L.). Mol. Breeding, 34(1): 159-166. DOI:10.1007/s11032-014-0026-2.

[88]. McInnis SM, Costa LM, Gutiérrez-Marcos JF, Henderson CA, Hiscock SJ (2005) Isolation and characterization of a polymorphic stigma-specific class III peroxidase gene from *Senecio squalidus* L. (Asteraceae). Plant Mol. Biol. 57(5): 659-677. DOI:10.1007/s11103-005-1426-9.

[89]. Horner HT and Palmer RG (1995) Mechanisms of genic male sterility. Crop Sci. 35(6): 1527-1535. DOI:10.2135/cropsci1995.0011183X003500060002x.

[90]. Chaudhury AM (1993) Nuclear genes controlling male fertility. Plant Cell 5: 1277–1283. DOI:10.1105/tpc.5.10.1277.

[91]. Yamaguchi T, Lee DY, Miyao A, Hirochika H, An G, Hirano HY (2006) Functional diversification of the two C-class MADS box genes OSMADS3 and OSMADS58 in *Oryza sativa*. Plant Cell, 18(1): 15-28. DOI:10.1105/tpc.105.037200.

[92]. De Neve C (2014) Male Sterility in Rice Pollen Development. Ghent University Library, Universiteit Gent. Faculteit Bio-ingenieurswetenschappen Vakgroep LA15.

Chapter IV

Toward a First High-quality Genome Draft for Marker-assisted Breeding in Leaf Chicory, Radicchio (*Cichorium intybus* L.)

Published: Galla G, Ghedina A, Tiozzo CS, Barcaccia G (2016) Toward a first high-quality genome draft for marker-assisted breeding in leaf chicory, Radicchio (*Cichorium intybus* L.). In Plant Genomics, Dr. Ibrokhim Y. Abdurakhmonov (Ed.), InTech. DOI:10.5772/61747.

Abstract

Radicchio (*Cichorium intybus* subsp. *intybus* var. *foliosum* L.) is one of the most important leaf chicories, used mainly as a component for fresh salads. Recently, we sequenced and annotated the first draft of the leaf chicory genome, as we believe it will have an extraordinary impact from both scientific and economic points of view. Indeed, the availability of the first genome sequence for this plant species will provide a powerful tool to be exploited in the identification of markers associated with or genes responsible for relevant agronomic traits, influencing crop productivity and product quality. The plant material used for the sequencing of the leaf chicory genome belongs to the Radicchio of the Chioggia type. Genomic DNA was used for library preparation with the TruSeq DNA Sample Preparation chemistry (Illumina). Sequencing reactions were performed with the Illumina platforms HiSeq and MySeq, and sequence reads were then assembled and annotated.

We are confident that our efforts will extend the current knowledge of the genome organization and gene composition of leaf chicory, which is crucial for developing new tools and diagnostic markers useful for our breeding strategies in Radicchio.

Keywords: Genome draft, marker-assisted breeding, gene prediction, SSR markers, SNP calling

Introduction

The common Italian name of Radicchio was adopted in recent years by all the most internationally used languages and indicates a highly differentiated group of chicories, with red or variegated leaves. Radicchio (*Cichorium intybus* subsp. *intybus* var. *foliosum* L.) is currently one of the most important leaf chicories, used mainly as a component for fresh salads but also very often cooked and prepared differently according to local traditions and alimentary habits [1].

This plant species belongs to the Asteraceae family and includes several cultivar groups whose commercial food products are the leaves, namely Witloof, Pain de sucre, and Catalogne, as well as several types of Radicchio.

From the reproductive point of view, Radicchio is prevalently allogamous, due to an efficient sporophytic self-incompatibility system, proterandry and gametophytic competition favouring allo-pollen grains and tubes [1]. Probably known by the Egyptians and used as food and/or medicinal plants by the ancient Greeks and Romans, this species gradually underwent a process of naturalization and domestication in Europe during the past few centuries. This plant has become part of both natural and agricultural environments of Italy. Currently, among the different biotypes of leaf chicories, the so-called Radicchio of Chioggia, native to and very extensively grown in northeastern Italy, is the Radicchio cultivar acquiring more and more commercial interest worldwide. In Italy, the Radicchio of Chioggia is cultivated on a total area of approximately 16–18,000 ha, half of which is in the Veneto region, with a total production of approximately 270,000 tons (more than 60% obtained using professional seeds), reaching an overall turnover of approximately € 10,000,000 per year.

Grown plant materials are usually represented by landraces or their directly derived synthetics that are known to possess a high variation and adaptation to the natural and anthropological environment where they originated from and are still cultivated. These populations are characterized by high-quality traits and have been maintained or even improved over the years by local farmers through phenotypical selection according to their own criteria and more recently by seed companies through genotypical selection following intercross or polycross schemes combined with progeny tests to obtain populations showing superior DUS scores for both agronomic and commercial traits. The breeding programs

currently underway by local firms and regional institutions exploit the best landraces and aim to isolate individuals amenable for use as parents for the constitution of narrow genetic base synthetic varieties and/or to select inbred lines suitable for the production of heterotic F1 hybrids [2]. In recent years, phenotypic evaluation trials are increasingly assisted by genotypic selection procedures through the use of molecular markers scattered throughout the genome. In fact, markerassisted breeding allows the identification of the parental individuals or the inbred lines showing the best general or specific combining ability in order to breed synthetics and hybrids, respectively.

Radicchio, like the other leaf chicories, is diploid ($2n=2x=18$) and is characterized by an estimated haploid genome size of approximately 1.3 Gb. In recent years, three distinct saturated molecular linkage maps were constructed for leaf chicories, covering approximately 1,200 cM [3-4-5]. Its linkage groups were mainly based on neutral SSR markers, but many ESTderived SNP markers were also mapped. A method for genotyping elite breeding stocks of Radicchio, both local and modern varieties, assaying mapped SSR marker loci possibly linked to EST-rich regions and scoring PIC>0.5, was recently developed using multiplex PCRs [6].

Here, we are dealing with a research and development project aimed at sequencing and annotating the first draft of the leaf chicory genome as we believe it will have an extraordinary impact from both scientific and economic points of view. Indeed, the availability of the first genome sequence for this plant species will provide a powerful tool to be exploited in the identification of markers associated with or genes responsible for relevant agronomic traits, influencing crop productivity and product quality. As an example, data and knowhow produced in this research project will be useful for detailed studies of the genetic control of male-sterility and self-incompatibility in this species. The plant material that we used for the sequencing of the leaf chicory genome belongs to the Radicchio of Chioggia type, specifically to the male-fertile inbred line named SEG111. This type was chosen as the most suitable accession based on the following criteria: i) the commercial relevance of the variety of origin; ii) the availability of clonal materials; iii) robust phenotypic and genotypic characterization; iv) a high degree of homozygosity (80%); and v) high breeding value as pollen parent of F1 hybrids. Sequencing reactions of the genomic DNA library were performed with Illumina HiSeq and MySeq platforms to

combine the high number of reads originated by the former with the longer sequences produced by the latter. Here, we report original data from the bioinformatic assembly of the first genome draft of Radicchio, along with the most relevant findings that emerged from an extensive *de novo* gene prediction and *in silico* functional annotation of more than 18,000 unigenes. Analyses were performed according to established computational biology protocols by taking advantage of the publically available reference transcriptome data for *Cichorium intybus* [7]. The main preliminary findings on the genome organization and gene composition of Radicchio are presented, and the potentials of newly annotated expressed sequences and diagnostic microsatellite markers in breeding programs are critically discussed.

Materials and methods

Plant materials

Plant materials used for the sequencing belong to a variety of commercial relevance of the Radicchio of Chioggia type. The clone chosen derives from the inbred line SEG111 and shows a degree of homozygosity equal to 80% [6]. In particular, this clone was obtained by several cycles of selfing from plants yearly selected on the basis of a robust phenotypic and genotypic characterization, being also characterized by high-quality agronomic traits on farm and the ability to be easily cloned in vitro.

DNA isolation and sequencing

DNA was isolated from 150 mg of fresh leaf tissue using a CTAB-based protocol [8]. The eventual contamination of RNA was avoided with an RNase A (Sigma-Aldrich) treatment. DNA samples were eluted in 80–100 μL of 0.1× TE buffer (100 mM Tris-HCl 1, 0.1 mM EDTA, pH=8). The integrity of the extracted DNA samples was estimated through electrophoresis in 0.8% agarose/1× TAE gels containing 1× SYBR Safe DNA Gel Stain (Life Technologies, USA). The purity and quantity of the DNA extracts were assessed with a NanoDrop spectrophotometer (Thermo Scientific, USA). Then, 1 μg of high-quality DNA was used for library preparation with the TruSeq DNA Sample Preparation chemistry (Illumina). Sequencing reactions were performed with the Illumina platforms: HiSeq (1 lane, 2 × 100 bp) and MySeq (1 lane, 2 × 300 bp).

De novo assembly and annotation

All high-quality reads generated from the two sequencing reactions were assembled in a single reference genome. Assemblies were attempted with three pieces of software: i) Velvet [9]; ii) SPAdes [10]; and iii) CLC Genomics Workbench 6.5 (Qiagen). The average coverage was estimated for the run HiSeq by calculating the frequency distribution of 25-mers [11].

To annotate all assembled contigs, a BLASTX-based approach was used to compare the *C. intybus* sequences to a subset of the NR protein collection that was made by focusing on the clade pentapetalae [12]. Moreover, the GI identifiers of the best BLASTX hits, having E-

value ≤1.0E-15 and similarity ≥70%, were mapped to the UniprotKB protein database [13] to extract Gene Ontology annotations [14] and KEGG terms [15] for functional annotations. Further enrichment of enzyme annotations was made with the BLAST2GO software v1.3.3 using the function “direct GO to Enzyme annotation”. The BLAST2GO software v1.3.3 [16-17] was used to reduce the complexity of the data and perform basic statistics on ontological annotations, as reported by Galla *et al.* [18].

SSRs were detected among the 522.301 contigs via MISA [19]. The parameters were adjusted to identify perfect and complex mono-, di-, tri-, tetra-, penta-, and hexanucleotide motifs with a minimum of 49, 13, 9, 8, 8, and 8 repeats, respectively. Repeated elements were detected with a BLASTN-based approach using a PGSB Repeat Element Database in all BLAST searches [20].

The parameters set for the identification of Transposable Elements (TEs) were: reward 1, penalty 1, gap_open 2, gap_extend 2, word_size 9, dust no. An E-value cutoff of 1.0E-9 was adopted to filter the BLAST results.

Two public *C. intybus* transcriptomes CHI-2418 and CHI-Witloof originally developed from

plant seedlings [7] corresponding to a wild accession of leaf chicory and a cultivated variety of witloof, respectively, were mapped to the reference genome using the CLC Genomics Workbench V7.02 (Qiagen). Mappings were performed with default mapping parameters, including mismatch cost: 2; insertion cost: 3; deletion cost: 3; length fraction: 0.5; and similarity fraction: 0.8. Non-specific matches were ignored and not included in the annotation tracks. For nucleotide variant analysis, the appropriate reference masking options were used to map transcriptome reads selectively over the sequences annotated as CDS or TEs. The variant detection analysis was done by using the Basic Variant Detection tool of the CLC Genomics Workbench V7.02 (Qiagen) with default parameters. As general filters, positions with coverage above 100,000 were not considered. Base quality filters were turned on and set to default parameters. All variants included in homopolymer regions with minimum length of 3nt, and with frequency below 0.8 were also removed from the dataset. As coverage and count filters, all variants with a minimum count lower than 20 were discarded.

Results

Genome assembly statistics

To obtain the first genome draft of leaf chicory, a single genomic library produced from the inbred line SEG111 was sequenced using the Illumina MySeq and HiSeq platforms. Here, we report the genome assembly results derived from the CLC Genomic Workbench assembly output. Figure 1 describes the frequency distribution of 25-mers in the HiSeq data.

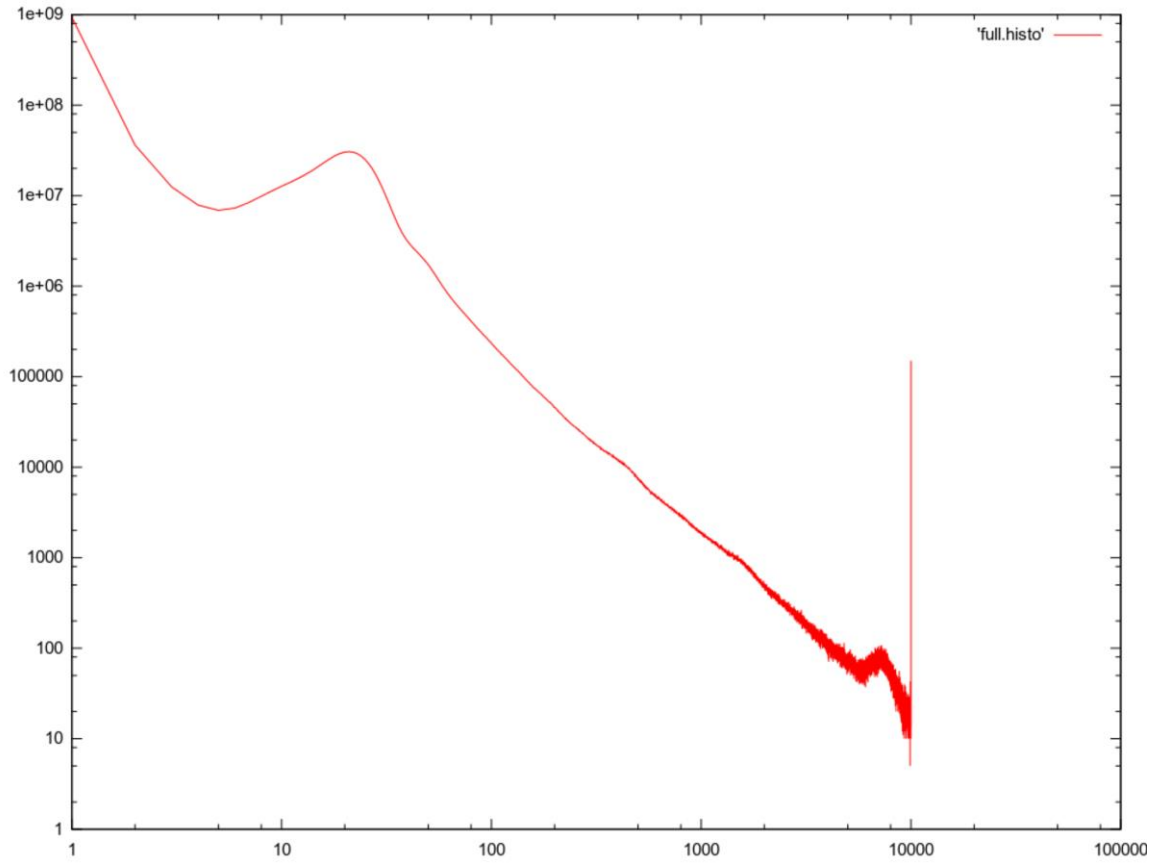


Figure 1. Frequency distribution of 25-mers in the HiSeq data (logarithmic scale for both axes)

The data shown suggest that the average coverage in the HiSeq run is approximately 21×. Additionally, the curve indicates that a certain number of sequences are present with a relatively high frequency within the genome. This might indicate that repeated elements are

relatively abundant within the genome. As a consequence, the estimated size of the assembled genome draft is 760 Mb.

We obtained 58,392,530 and 389,385,400 raw reads through the MySeq and HiSeq platforms, respectively. The *de novo* assembly of the two datasets in a unique reference genome draft assembled 724,009,424 nucleotides into 522,301 contigs (**Table 1**). The maximum contig length was equal to 379,698 bp, whereas the minimum contig length was set to 200 bp, with an average contig length of 1,386 bp. Overall statistics are summarized in **Table 1.**

Table 1. Summary statistics of the sequence assembly generated from *Cichorium intybus*. The length distribution of the contig size, expressed in base pairs, is reported in **Figure 2.**

Total number of contigs	522,301
Total No. of assembled nucleotides (nt)	724,009,424
GC percentage	34.8%
Average contig lenght (bp)	1,386
Minimum contig length (bp)	200
Maximum contig length (bp)	379,698
N75	1,051
N50	3,131

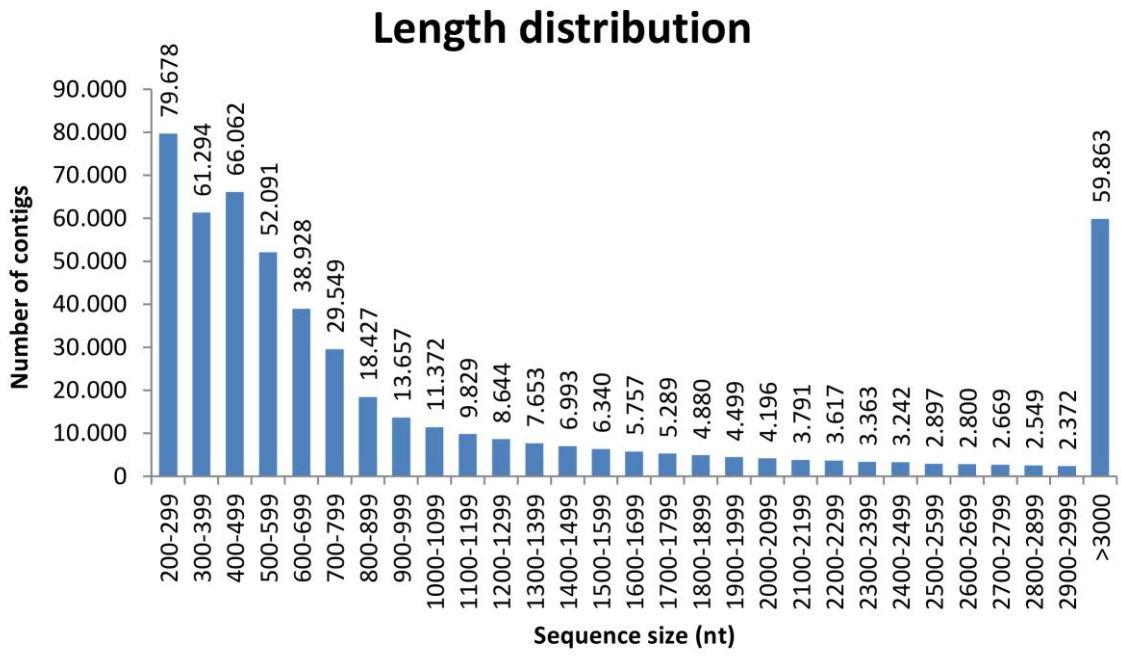


Figure 2. Distribution of length of contigs for leaf chicory

As much as 68.9% of the recovered sequences are contained within a length spanning from 200 nt to 999 nt. The interval length ranging between 1,000 nt and 2,999 nt is represented by 19.7% of the assembled contigs, whereas the proportion of contigs whose length is higher or equal to 3,000 nt corresponds to 11.5%.

We searched the genome sequence assembly for TEs and estimated their abundance using a BLASTN strategy. The proportion of base pairs annotated as TEs out of the total amount of assembled nucleotides was equal to 6.3% (**Table 2**).

Table 2. Classification statistics of transposable elements (TEs) in Radicchio genome draft assembly.

Key	Classification	Number	Percentage (%)	Length (bp)	Percentage over the assembled genome
02.01	Class I retroelement	273	0.19	85,241	0.012%
02.01.01	LTR Retrotransposon	82,260	56.55	19,658,874	2.715%
02.01.01.05	Ty1/copia	35,802	24.61	17,519,102	2.420%
02.01.01.10	Ty3/gypsy	23,651	16.26	7,121,605	0.984%
02.01.02	non-LTR Retrotransposon	354	0.24	106,259	0.015%
02.05	Class II: DNA Transposon	1,976	1.36	713,119	0.098%
02	Unclassified mobile element	861	0.59	199,301	0.028%
10 / 90 / 99	High Copy Number Genes and Aditional attributes	283	0.19	51,577	0.007%
Total		145,462	100.0	45,455,078	6.278%

The retroelements were the most abundant elements (>97% of the total). Within the major class of retroelements, Long Terminal Repeat (LTR) retrotransposons proved to be the dominant class (56.55%) in the leaf chicory genome. Moreover, the Copia-type (24.61%) and the Gypsy-type (16.26%) appeared to be the most abundant LTR retrotransposons. A total of 273 (0.2%) elements were annotated as retroelements, but they lacked the assignation to a specific class based on sequence similarity and conservation. Non-LTR retrotransposons were detected to a very low extent (0.24%). Less than 2% of the total repeat elements were annotated as DNA transposons.

Discovery of SSR loci

Overall, we identified 66,785 SSR containing regions. As many as 52,186 and 11,501 sequences proved to contain one or more microsatellites, respectively. These numbers included 1,226 mononucleotide SSR motifs (which were no longer taken into account for further computations).

We found a total number of di- or multinucleotide SSR motifs equaling 65,559. The most common SSR elements were those showing a dinucleotide motif (89.0%), followed by trinucleotide (7.1%) and tetranucleotide (3.0%) ones. Microsatellites revealing a pentanucleotide and hexanucleotide motif were less than 1.0% of the total. Overall data are summarized in **Table 3**.

Table 3. Number of SSRs detected in the Radicchio genome draft assembly. For each type of motif, the number of SSRsidentified in the range of repeated numbers is reported. Albeit present in the genome, mono-nucleotide SSRs were notconsidered in this analysis.

Type of motif	Range of repeat numbers				Total	Percentage (%)
	8-12	13-17	18-22	>22		
Di-nucleotide	0	8,333	7,100	42,913	58,346	89.0
Tri-nucleotide	1,822	1,769	762	321	4,674	7.1
Tetra-nucleotide	1,114	475	205	202	1,996	3.0
Penta-nucleotide	69	23	0	2	94	0.1
Hexa-nucleotide	359	80	8	2	449	0.7
Total	3,364	10,680	8,075	43,440		
Percentage (%)	5.1	16.3	12.3	66.3		

Functional annotation of contig sequences

Functional annotation of the assembled contigs was performed with a BLASTX approach, according to which all contig sequences were used to query different public protein databases (**Table 4**).

Table 4. Summary statistics of functional annotations for leaf chicory genome sequences in public protein databases.As for the NR database, only the protein sequences from the clade pentapetalae of eudicots were considered. The Arabidopsis proteome used in all BLAST analysis was TAIR10.

Public database	Number of Hits (gene models)	Number of *C. intybus* contigs
NR	38,782	80,862
Arabidopsis	16,689	50,417
GO	14,073	45,381
KEGG	4,512	22,273

The database enclosing all public protein sequences belonging to the pentapetalae clade of the eudicots, which includes the sub-clades of rosids and asterids to which leaf chicory belongs, provided a total of 38,782 hits. The proteome of *Arabidopsis thaliana* alone scored 16,689 hits when an E-value cutoff of 1.0E-15 was applied for the screening of the most reliable BLASTX hits.

Two public *C. intybus* transcriptomes originally developed from plant seedlings and provided by UC DAVIS, the Compositae Genome Project (CHI-2418 and CHI-Witloof) [7] were mapped to the reference genome using the appropriate mapping function of the CLC Genomics Workbench.

By doing so, we were able to map 76.5% and 78.0% of the sequences, respectively. Data derived from the mapping of two *C. intybus* transcriptomes were used to integrate the annotation of the assembled contigs. BLAST and mapping data integration increased the BLAST-based annotation with an additional set of 1,995 contigs.

Arabidopsis matches were used to retrieve both GO and KEGG annotations from public databases. We could finally assign one or multiple GO terms to 45,381 leaf chicory genome contigs. The analysis performed against the GO illustrate 14,073 genes annotated with terms belonging to one or multiple vocabularies. Of these, 24,634 contigs were annotated for their putative biological process, 39,118 contigs were related to a molecular function, and 37,561 contigs were associated to a specific cellular component. **Figure 3** shows the fine distribution of the 14,073 hits caught by our Radicchio contigs from the TAIR database according to the aforementioned three GO categories.

Among all the terms underlined by the GO vocabulary for the biological process, our investigations were focused on terms related to the response to biotic and abiotic stresses (**Figure 4**), hormonal responses (**Figure 5**), and flower and seed development (**Figure 6**). Of the 15 most interesting processes for molecular breeding in leaf chicory, 7 and 8 were linked to biotic and abiotic stresses, respectively (see **Figure 4**). The ontological terms were assigned to 2,388 and 3,844 genome contigs, respectively.

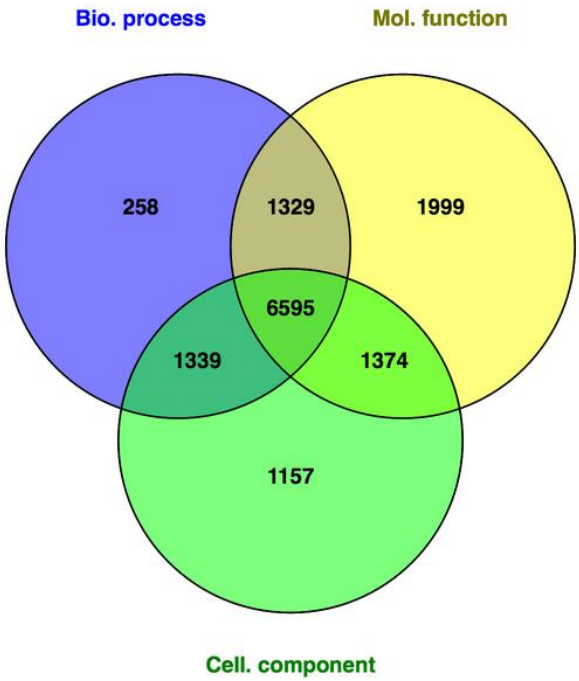


Figure 3. Venn diagram showing the fine distribution according to GO terms of the 14,073 *A. thaliana* hits matching our leaf chicory contigs

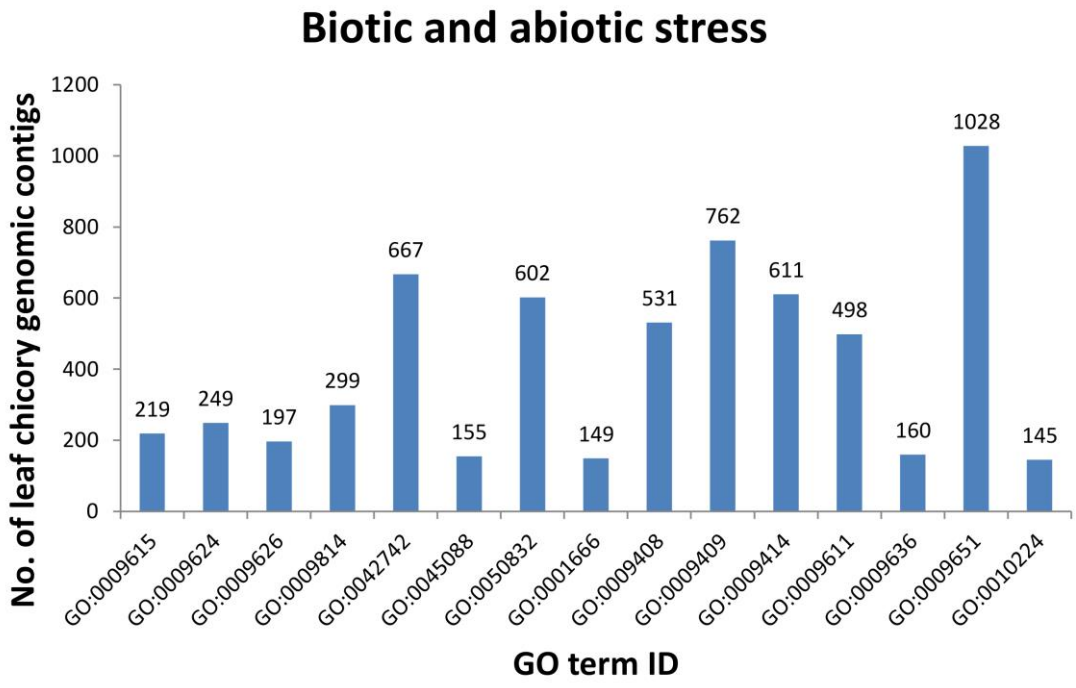


Figure 4. Number of *C. intybus* genomic contigs for response to biotic (the first 7) and abiotic (the last 8) stress.

The computational analysis for the identification of SSR elements within these contigs unveiled 495 motifs linked to biotic stresses and 841 motifs associated with abiotic stresses. Among the biotic stresses, the most abundant gene ontology (GO) term was GO:0042742, which corresponds to the "defense response to bacterium" and shows a match with 667 genome contigs containing 135 microsatellites. Concerning the abiotic stresses, the GO term assigned with the higher frequency was GO:0009651, which accounts for processes related to "response to salt stress" and matches 1,028 genome contigs containing 249 microsatellites.

Data of hormonal responses and processes of flower and seed development are reported in **Figures 5** and **6**. The analysis for hormonal responses noted nine different GO terms, for a total of 3,344 genome contigs, and 833 SSR elements linked to these sequences and terms. In particular, the term "response to jasmonic acid stimulus" (GO:0009753) was the most represented, with 478 matches with different genome contigs, including 118 SSR motifs (**Figure 5**).

Results of the GO term annotation of genome contigs according to the flower and seed developmental processes are reported in **Figure 6**.

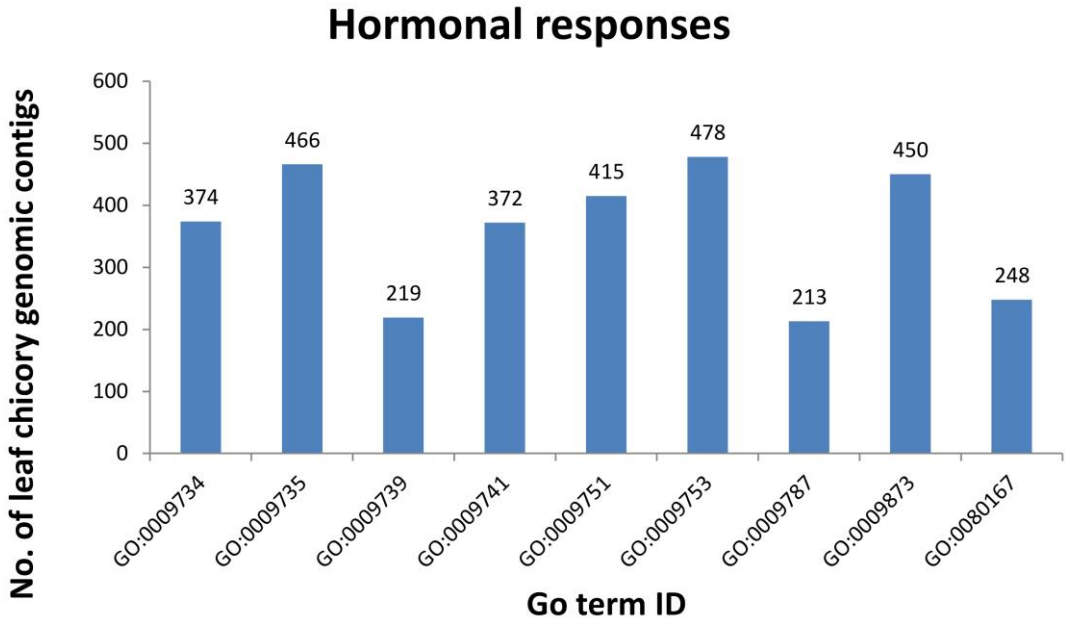


Figure 5. Number of C. intybus genomic contigs for hormonal response

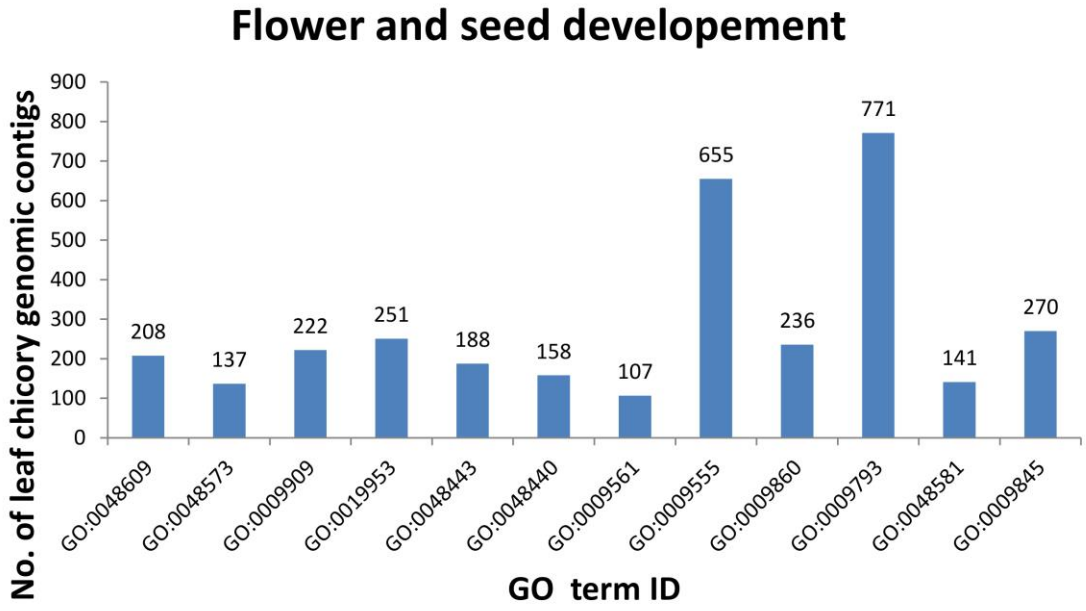


Figure 6. Number of *C. intybus* genomic contigs for flower and seed (only the last 3) development

The flower development process was embraced by selecting nine ontological terms, whereas three terms were assigned to seed development and seed germination. A total of 2,162 contigs were annotated with GO terms related to flower development; 496 of these were also annotated for the presence of one or multiple SSRs. In particular, the term "pollen development" (GO: 0009555) was the most abundant, with 655 contigs containing 153 SSR motifs.

As far as the seed development process is concerned, we annotated 1,182 contigs linked to this GO term, 273 of which co-localized with one or multiple SSRs. Among these, the most abundant ontological term was "embryo development ending in seed dormancy" (GO: 0009793) as it is assigned to 771 contigs, co-localizing with 171 SSR elements.

Using the Kyoto Encyclopaedia of Genes and Genomes database (http://www.genome.jp/kegg/), a total of 22,273 contigs enabled the mapping of 795 enzymes to 157 metabolic pathways.

Among the metabolic pathways with the highest number of mapped reads, we found fructose and mannose metabolism (418 gene models matched), phenylpropanoid biosynthesis (415 gene models matched) and tryptophan metabolism (380 gene models matched). The biosynthetic pathway of flavonoid biosynthesis, described in map:00941, is

relevant as the biosynthesis of flavonoid is directly connected to the synthesis of anthocyanin (**Figure 7**), whose accumulation contributes to the pigmentation of leaf chicories. This map includes 236 gene models that were assigned to 14 unique enzymes, including CHS (CHALCONE SYNTHASE), CHI (CHALCONE ISOMERASE), and ANS (ANTHOCYANIDIN SYNTHASE), among others.

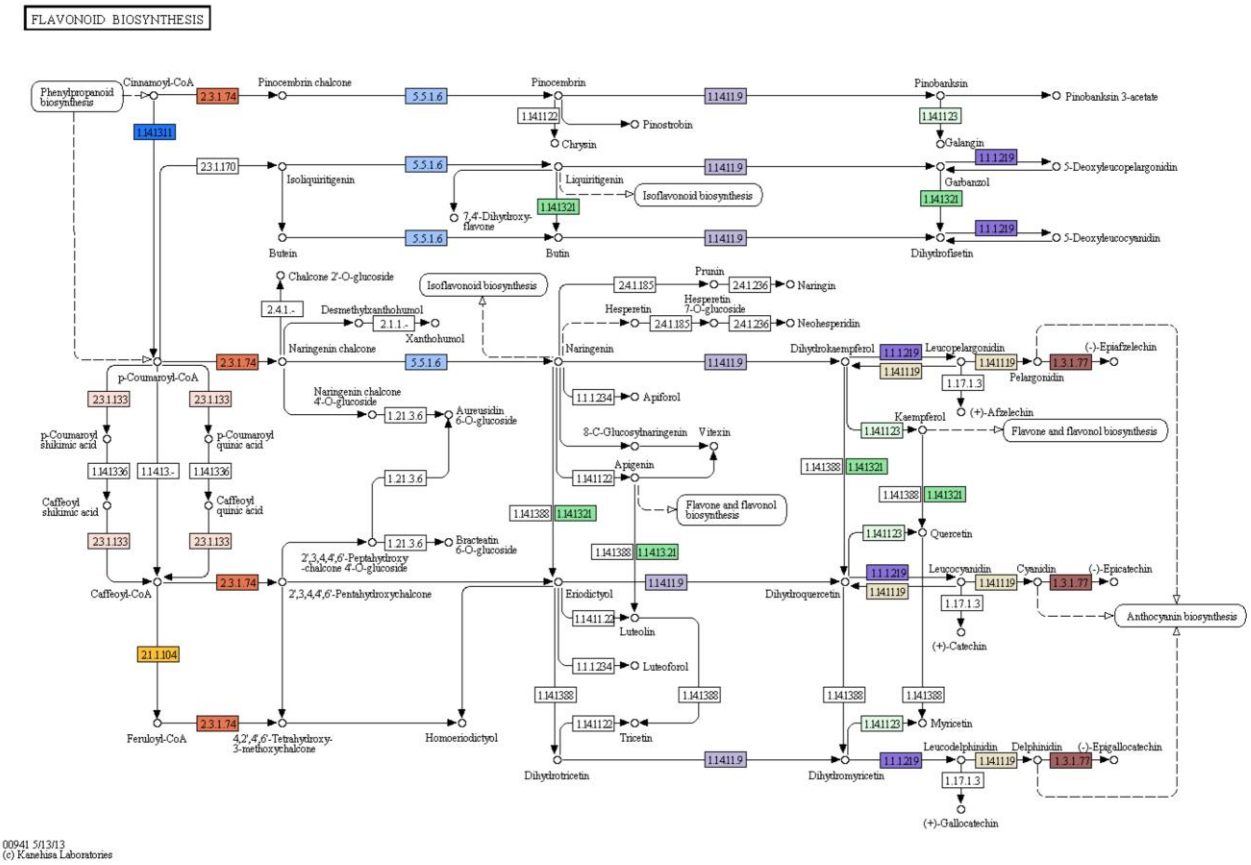


Figure 7. KEGG pathway for flavonoid biosynthesis (Map:00941)

KEGG data related to a number of selected metabolic pathways were exploited to find SSR regions potentially associated with highly valuable phenotypes in this plant species. The number of SSRs putatively linked to the most interesting phenotypic traits with breeding values in leaf chicory is displayed in **Table 5**.

Table 5. Number of SSRs located in contig sequences annotated for the presence of proteins with known enzymatic activity in relevant metabolic pathways for the breeding of leaf chicory.

KEGG map ID	Metabolic pathway	Characteristic	No. of SSRs
map00909	Sesquiterpenoid and triterpenoid biosynthesis	Bitter taste	107
map00053	Ascorbate and aldarate metabolism	Vitamin C content	172
map00940	Phenylpropanoid biosynthesis	Leaf color	281
map00941	Flavonoid biosynthesis	Leaf color	173
map00942	Anthocyanin biosynthesis	Leaf color	180
map00943	Isoflavonoid biosynthesis	Leaf color	5
map00944	Flavone and flavonol biosynthesis	Leaf color	128
map00040	Pentose and glucuronate interconversions	Response to cold	96
map00051	Fructose and mannose metabolism	Response to cold	259
map00052	Galactose metabolism	Response to cold	31
map00061	Fatty acid biosynthesis	Response to cold	39
map00260	Glycine, serine and threonine metabolism	Response to cold	60
map00290	Valine, leucine and isoleucine biosynthesis	Response to cold	13
map00330	Arginine and proline metabolism	Response to cold	55
map00410	beta-Alanine metabolism	Response to cold	16
map00480	Glutathione metabolism	Response to cold	48
map00500	Starch and sucrosa metabolism	Response to cold	164
map00561	Glycerolipid metabolism	Response to cold	159
map00564	Glycerophospholipid metabolism	Response to cold	124
map00592	alpha-Linolenic acid metabolism	Response to cold	66
map00710	Calvin cycle	Response to cold	28
map00780	Biotin metabolism	Response to cold	18
map00960	Tropane, piperidine and pyridine alkaloid biosynthesis	Response to cold	97

Considering the overall grouping of selected metabolic pathways, we identified many microsatellite sequences putatively linked to important traits, according to their potential effect on plant characteristics. For instance, 107 SSRs were linked to bitter taste, 172 SSRs were associated with vitamin C biosynthesis and metabolism, and 767 SSRs located in sequence contigs encoding enzymes of the flavonoid and anthocyanin biosynthetic

pathways, thus potentially associated with the leaf color. The most represented characteristic is the response to cold. For this trait, we analyzed 16 different metabolic pathways that altogether led to the selection of 1,273 microsatellites potentially associated with one or multiple genes actively involved in the plant response to cold eventually, but not exclusively, through the accumulation of sugar.

We also performed the calling of nucleotide variants. Stringent quality criteria were used for discriminating sequence variations from sequencing errors and mutations introduced during cDNA synthesis. Only sequence variations with mapping quality scores over the established thresholds were annotated, leading to the identification of 123,943 and 121,086 variants that were present only in the leaf chicory transcriptome CHI-2418 (wild type) or the Witloof transcriptome CHI-Witloof (cultivated type), respectively. A total of 119,729 variants were shared by both *C. intybus* transcriptomes. The average number of variants per contig ranged from 9.5 to 10.5 in the two assemblies (**Table 6**), yielding one single variation per 100 bp in both cases.

Table 6. Summary statistics of nucleotide variants restricted to gemonic regions of Radicchio annotated as CDS and Transposable Elements. Nucleotide variants were detected by using the transcriptomes CHI-2418 (wild type leaf chicory) and CHI-Witloof (cultivated Witloof type). For each transcriptome, the number of contigs displaying one or multiple variants, the number of variants and the number of variants per contigs are indicated. The number of variants per 100bp is also reported. Variants present in both transcriptomes are indicated as shared.

	CDS - 29175 contigs			TEs - 122745 contigs		
	CHI-2418	**CHI-Witloof**	**Comm.**	**CHI-2418**	**CHI-Witloof**	**Comm.**
Number of contigs	12,725	12,739	11,419	2,016	1,924	1,554
Number of variants	123,843	121,086	119,729	10,662	10,651	10,246
Number of variants/contigs	9.75	9.52	10.52	5.29	5.54	6.61
Number of variants/100bp	0.99 (1.14)	0.98 (1.14)	1.14 (2.05)	3.26 (3.64)	3.16 (3.50)	5.42 (8.88)
SNV	115,678	113,049	107,255	9,532	9,605	9,006
MNV	5,367	5,439	8,475	507	441	261
Insertion	2,044	2,036	2,166	556	552	714
Deletion	754	562	1,833	67	53	265

The vast majority of variants were Single Nucleotide Variants (SNVs), whereas Multi Nucleotide Variants (MNVs), Insertions, and Deletions were found to a considerably lower extent (**Table 6**). On average, the proportion of SNVs and MNVs was comparable in the CDS and TE contigs and equal to about 90% and 5%, respectively.

Among all contigs annotated as TEs, those characterized by the presence of one or multiple variants were 10,662 and 10,651 for the two transcriptomes (**Table 6**). The average number of variants per contig was equal to 5.3 and 5.5. Despite the relatively low abundance of polymorphic residues in these regions, the average number of variants per 100 bp was equal to 3.3 and 3.2. Single Nucleotide Polymorphisms (SNPs) were by far the most abundant type of variants in TEs as well as in CDS regions (**Table 6**). In particular, transversions and transitions were on average 37% (ranging from 35.6% and 37.8%) and 63% (ranging from 62.2% and 64.4%) of the point mutations, respectively. The total number of nonsynonymous SNPs calculated with the reference transcriptomes was equal to 13,559 (10.9%) and 11,197 (9.2%) for wild-type leaf chicory and cultivated Witloof accessions, respectively.

Discussion

Here, we report the uncovering of the first draft of the Radicchio genome. This highly relevant discovery was achieved by combining the recent advancement of next-generation sequencing technologies on the public side with the significant investment of financial resources in research and development on the private side.

Currently, conventional agronomic-based selection methods are supported by molecular marker-assisted breeding schemes. In recent years, we have demonstrated that the constitution of F1 hybrids is not only feasible in a small experimental scheme but also realizable and profitable on a large commercial scale (*e.g.*, registered CPVO varieties TT4070/F1, TT5010/F1, TT5070/F1, and TT4010/F1 in progress). F1 hybrids are varieties manifesting heterosis, or hybrid vigor, which refers to the phenomenon in which highly heterozygous progeny plants obtained by crossing genetically divergent inbred or pure lines exhibit greater biomass, faster speed of development, higher resistance to pests and better adaptation to environmental stresses than the two homozygous parents. Critical steps of an applicative breeding program are the production of parental inbreds. Two highly relevant factors in this context are the selection of self-compatible genotypes, to be used as pollen donors, and the identification of male-sterile genotypes, to be used as seed parents in large-scale crosses [21-22].

It is worth mentioning that there are several reasons why the constitution of F1 hybrids is a strategic choice for a seed company. First, the crop yield of modern F1 hybrid varieties is usually much higher than that of traditional OP or synthetic varieties. Second, the uniformity of F1 populations and the way to legally protect their parental lines allow a seed company to adopt a plant breeder's rights, promoting genetic research and development programs that are very expensive and require many years. Finally, the need for breeding hybrid varieties also promotes the preservation of local varieties because the selection of appropriate inbred or pure lines as parents in pairwise cross-combinations requires the exploration and exploitation of germplasm resources. Our expectation is that F1 hybrid varieties will be bred and adopted with increasing frequency in Radicchio. Consequently, we invested in the sequencing and annotation of the first draft of the leaf chicory genome as it will have an extraordinary impact from both scientific and economic points of view. Indeed, the availability of the first genome sequence for this plant species will provide a

powerful tool to be exploited in the identification of markers associated with or genes responsible for relevant agronomic traits, influencing crop productivity and product quality. As an example, data and knowhow produced in this research project will be capitalized on in subsequent years to plan and develop basic studies and applied research on male-sterility and self-incompatibility in this species.

The availability of high-quality sequencing platforms (*i.e.*, Illumina) on the one hand, and specific and high-performing software for genome data assembly and gene set analysis on the other, made this project feasible. High-quality genomic DNA libraries were used for sequencing reactions performed with the Illumina platforms HiSeq and MySeq, originating a total of 197 million (mln) short reads and 29 mln longer sequences passing quality filters, respectively, which were then bioinformatically assembled to obtain the first genome draft. On the basis of this strategy, the genome draft of leaf chicory is composed of approximately 500,000 contigs, forming approximately 720 Mb. Based on the distribution of 25-mer frequencies, we estimated that the genome coverage is close to 25X. The same distribution also indicates that a significant part of the genome might be composed of highly repeated elements, as indicated by the number of k-mers that appears to be present with high frequency.

Nucleotide variant calling for the Radicchio genome showed comparable number of polymorphisms in the pairwise comparisons with the two publically available transcriptomes, originally developed from seedlings of two leaf chicory accessions (*i.e.*, wild and cultivated types). The total number of variants discovered in the CDS regions was shown to be approximately 10 times higher than the ones found in the TEs. This result might be a consequence of low expression, or silencing, of numerous transposable elements at the level of plant seedlings, as indicated by the finding that the mapping of the two transcriptomes to the reference genome failed to align sequences to about 98% of the contigs annotated as TEs. Noteworthy, the number of variations per 100 base pairs was significantly higher in the TEs than in the CSD sequences. This result might be explained by the accumulation of mutations in noncoding sequences, as most of the TEs are.

Overall, Single Nucleotide Variants (SNVs) were the most common variants compared with In/Del mutations. Since SNP mutations very often result in silent mutations, their high proportion in the CDS regions was an expected result. In/Del mutations that usually occur

in silenced or functionally disrupted genes, along with noncoding regions, were found at a low rate in CDS regions.

TEs were found to occur, at least in one copy, in the 23.50% of the 522,301 contigs that constitute our chicory genome draft assembly. Retrotransposons proved to be the most abundant elements in the Radicchio genome. This finding is in agreement with data from previous studies [23-24-25-26]. It is worth mentioning that Copia-type elements were more abundant than Gypsy-type elements, forming the predominant subclass of LTR retrotransposons.

Although the amount of TEs of the totally assembled sequences was much lower than that reported for other species, the class ratio of the TE types corresponds to that found in previous studies [23-24-25-26]. Our estimate of TEs in leaf chicory is equal to 6.28% of the contigs length, which is much lower than amounts reported for soybean (59%), pigeonpea (52%), alfalfa (27%), trefoil (34%), and chickpea (40%) [25-27-28-29-30]. One of the reasons could be that our BLAST strategy chosen to find repeated elements in the genome was less efficient than specific software (*e.g.*, RepeatScout and RepeatMasker [31-32]). Another reason could be the lack of TEs in the assembled portion of the Radicchio genome due to the low complexity of these repeated DNA regions.

The BLAST strategy with the nonredundant (NR) pentapetalae protein database produced the best output in terms of similarity with our contigs. This is undoubtedly due to the availability of large collections of sequences from species taxonomically related to leaf chicory, such as *Beta vulgaris*, *Helianthus annuus,* and *Lactuca sativa*, among others. Unfortunately, the depth of annotation of these recently sequenced genomes is frequently not comparable to that of the long-studied *Arabidopsis thaliana*. Although BLAST results obtained by querying the NR database proved to be highly informative in terms of the number of hits producing alignments with significant E-value, the annotation of the leaf chicory assembled contigs was more successful when the *A. thaliana* database was used alone. Therefore, a possible alternative for future enrichment of the current annotation state would imply the use of software (*e.g.*, Blast2GO) that could extract the annotation codes from multiple BLAST hits, provide the appropriate specificity cutoff, and assign the mapped GO terms to the original query.

Our choice to use the TAIR10 database to annotate our sequence contigs led to the annotation of a large number of assembled sequences and provided precious information concerning the putative process, or eventually, the metabolic pathways in which genes are putatively active.

The ability to annotate a certain number of sequences is not only exclusively dictated by the length and quality of the query sequences but also by their match with orthologous sequences that need to be annotated in depth.

This would be the case of annotations for metabolic pathways not actively studied or present in *A. thaliana* and for processes whose study is hampered by biological or physical circumstances.

This might explain some discrepancies in annotations for male and female gametogenesis (**Figure 6**). From the graph, it is easy to understand the large discrepancy between the number of contigs presented for the term "Megagametogenesis" (GO:0009561), just 107, and the term "Pollen development" (GO:0009555), cited in the results as the most prevalent (more than six times that of megagametogenesis). We can suppose that this difference might not be due to a real difference in the number of genes involved in these two reproductive processes but rather to the lower number of genes known to be involved in female sporogenesis and gametogenesis.

Similarly, enzymes involved in the biosynthesis of germacren-type sesquiterpenoids, such as the germacrene-A synthase (EC:4.2.3.23), which are responsible for the biosynthesis of lactones associated with bitter taste in leaf chicory, are not known or properly characterized in *A.thaliana*.

Another fundamental finding of our study is the large number of SSR markers that were found in the assembled contigs. We can affirm that the leaf chicory genome shows an unexpected number and distribution of repeated sequences. Submitting our Radicchio draft to MISA software, we were able to reveal such a number of potential SSR markers. It is therefore interesting that we were able to link a reasonably large number of microsatellites to each item here presented for both GO terms and KEGG maps. In the results, we presented only a small selection of important characteristics that could be utilized in markerassisted selection and breeding programs in Radicchio. Together with SSRs, thousands of sequences that could be used in Single Nucleotide Polymorphism (SNP)

analysis were associated to fundamental biosynthetic pathways or metabolism enzymes. This is a crucial starting point for modern breeding in leaf chicories. It is noteworthy that further studies must be conducted to determine whether and how these potential markers could be exploited in molecular breeding programs. As a final step, gene prediction and annotation were also performed according to established computational biology protocols by taking advantage of the reference transcriptome data publically available for *Cichorium intybus* L. These sequences allowed us to learn the number, sequence, and role of the ~25.000 genes of the Radicchio's genome. This finding represents an important achievement for Italian agriculture genetics as a whole and opens new perspectives in both basic and applied research programs in Radicchio. It will have great impacts, potentials, and advantages in terms of breeding methods and tools useful for the constitution and protection of new varieties. Information obtained by the sequencing of the genome will be exploitable to detect and dissect the chromosomal regions where the genetic factors that control the expression of important agronomic and qualitative traits are located in Radicchio.

Modern marker-assisted breeding (MAB) technology based on traditional methods using molecular markers such as SSRs and SNPs, without relations to genetic modification (GM) techniques, will now be planned and adopted for breeding of vigorous and uniform F1 hybrids combining quality, uniformity, and productivity traits in the same genotypes.

In conclusion, our study will contribute to increase and reinforce the reliability of Italian seed firms and local activities of the Veneto region associated with the cultivation and commercialization of Radicchio plant varieties and food products; the seed market of this species will have the chance to become highly professional and more competitive at the national and international levels. To uncover the sequence of a given genome means to gain a robust scientific background and technological knowhow, which in short time can play a crucial role in addressing and solving issues related to the cultivation and protection of modern Radicchio varieties. In fact, we are confident that our efforts will extend the current knowledge of the genome organization and gene composition of leaf chicories, which is crucial in the development of new tools and diagnostic markers useful for our breeding strategies, and allow researchers for more focused studies on chromosome regions controlling relevant agronomic traits of Radicchio. In addition, conducting novel research

programs for the preservation and valorization of the biodiversity, still present in the Radicchio germplasm of the Veneto region, is very important and accomplished through the genetic characterization of the most locally dominant and historically important landraces using sequenced genome information of Radicchio presented in this work.

References

[1]. Lucchin M, Varotto S, Barcaccia G, Parrini P (2008) Chicory and Endive. In Handbook of Plant Breeding, Vegetables I: Asteraceae, Brassicaceae, Chenopodicaceae. Springer Science. New York. p. 1-46. DOI:10.1007/978-0-387-30443-4_1.

[2]. Barcaccia G, Pallottini L, Soattin M, Lazzarin R, Parrini P, Lucchin M (2003) Genomic DNA fingerprints as a tool for identifying cultivated types of radicchio (*Cichorium intybus* L.) from Veneto, Italy. Plant Breed. 122: 178-183. DOI:10.1046/j.1439-0523.2003.00786.x.

[3]. Cadalen T, Mörchen M, Blassiau C, Clabaut A, Scheer I, Hilbert J-L *et al*., Quillet MC (2010) Development of SSR markers and construction of a consensus genetic map for chicory (*Cichorium intybus* L.). Mol. Breeding. 25: 699-722. DOI:10.1007/s11032-009-9369-5.

[4]. Gonthier L, Blassiau C, Mörchen M, Cadalen T, Poiret M, Hendriks T, Quillet MC (2013) High-density genetic maps for loci involved in nuclear male sterility (NMS1) and sporophytic self-incompatibility (S-locus) in chicory (*Cichorium intybus* L., Asteraceae). Theor Appl Genet. 126(8), 2103-2121. DOI:10.1007/s00122-013-2122-9.

[5]. Muys C, Thienpont CN, Dauchot N, Maudoux O, Draye X, Van Cutsem P (2014) Integration of AFLPs, SSRs and SNPs markers into a new genetic map of industrial chicory (*Cichorium intybus* L. var. *sativum*). Plant Breed. 133(1): 130-137. DOI:10.1111/pbr.12113.

[6]. Ghedina A, Galla G, Cadalen T, Hilbert JL, Tiozzo CS, Barcaccia G (2015) A method for genotyping elite breeding stocks of leaf chicory (*Cichorium intybus* L.). BMC Res. Notes. 8(1): 1-12. DOI:10.1186/s13104-015-1819-z.

[7]. UC Davis. The Compositae Genome Project. Available from: http://compgenomics.ucdavis.edu/cgp_wd_assemblies.php#13427 [Accessed: June 2015].

[8]. Doyle JJ and Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 19: 11-15.

[9]. Zerbino DR, Birney E (2008) Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18(5): 821-829. DOI:10.1101/gr.074492.107

[10]. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19(5): 455-477. DOI:10.1089/cmb.2012.0021

[11]. Liu B, Shi Y, Yuan J, Hu X, Zhang H, Li N, Li Z, Chen Y, Mu D, FanW (2013) Estimation of genomic characteristics by analyzing kmer. Cornell University Library. Available from: http://arxiv.org/abs/1308.2012

[12]. NCBI. Available from: http://www.ncbi.nlm.nih.gov/

[13]. UniProt. Available from: http://www.uniprot.org/

[14]. Gene Ontology Consortium. [Updated: http://geneontology.org/].

[15]. KEGG: Kyoto Encyclopedia of Genes and Genomes. Available from: http://www.genome.jp/kegg/

[16]. Blast2GO. Available from: https://www.blast2go.com/

[17]. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 21(18): 3674-3676. DOI:10.1093/bioinformatics/bti610.

[18]. Galla G, Barcaccia G, Ramina A, Collani S, Alagna F, Baldoni L, Cultrera N. GM, Martinelli F, Sebastiani L, Tonutti P (2009) Computational annotation of genes differentially expressed along olive fruit development. BMC Plant Biol. 9(128). DOI:10.1186/1471-2229-9-128.

[19]. MISA - MIcroSAtellite identification tool [Internet]. [Updated: 5/14/02]. Available from: http://pgrc.ipk-gatersleben.de/misa/

[20]. Helmholtz Zentrum München - Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH). PGSB Plant Genome and Systems Biology [Internet]. Available from: http://pgsb.helmholtz-muenchen.de/plant/recat/

[21]. Barcaccia G and Tiozzo CS (2012) New male sterile Cichorium spp. mutant, parts or derivatives, where male sterility is due to a recessive nuclear mutation linked to a polymorphic molecular marker, useful for producing F1 hybrids of *Cichorium* spp. EU PATENT No. WO2012163389-A1. DOI: 1013140/2.1.4749.5044

[22]. Barcaccia G and Tiozzo CS (2014) New male sterile mutant of leaf chicory, including Radicchio, used to produce chicory plants and seeds with traits such as male sterility exhibiting cytological phenotype with shapeless, small and shrunken microspores in dehiscent anthers. USA PATENT No. US20140157448-A1. DOI:10.13140/2.1.3176.6400

[23]. Chan AP, Crabtree J, Zhao Q, Lorenzi H, Orvis J, Puiu D, Melake-Berhan A, Jones KM, Redman J, Chen G, Cahoon EB, Gedil M, Stanke M, Haas BJ, Wortman JR, Fraser-Liggett CM, Ravel J, Rabinowicz PD (2010) Draft genome sequence of the oilseed species *Ricinus communis*. Natur Biotechnol. 28: 951-956. DOI:doi:10.1038/nbt.1674

[24]. Rahman AYA, Usharraj AO, Misra BB, Thottathil GP, Jayasekaran K, Feng Y, Hou S, Ong SY, Ng FL, Lee LS, Tan HS, Sakaff MKLM, Teh BS, Khoo BF, Badai SS, Aziz NA, Yuryev A, Knudsen B, Dionne-Laporte A, Mchunu NP, Yu Q, Langston BJ, Freitas TAK, Young AG, Chen R, Wang L, Najimudin N, Saito JA, Alam M (2013) Draft genome sequence of the rubber tree *Hevea brasiliensis*. BMC Genomics. 14(75). DOI:10.1186/1471-2164-14-75.

[25]. Jain M, Misra G, Patel RG, Priya P, Jhanwar S, Khan AW, Shah N, Singh VK, Garg R, Jeena G, Yadav M, Kant C, Sharma P, Yadav G, Bhatia S, Tyagi AK, Chattopadhyay D (2013) A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.). Plant J. 74: 715-729. DOI:10.1111/tpj.12173.

[26]. He N, Zhang C, Qi X, Zhao S, Tao Y, Yang G, Lee T-H, Wang X, Cai Q, Li D, Lu M, Liao S, Luo G,He R, Tan X, Xu Y, Li T, Zhao A, Jia L, Fu Q, Zeng Q, Gao C, Ma B, Liang J, Wang X, Shang J, Song P, Wu H, Fan L, Wang Q, Shuai Q, Zhu J, Wei C, Zhu-Salzman K, Jin D, Wang J, Liu T, Yu M, Tang C, Wang Z, Dai F, Chen J, Liu J, Zhao S, Lin T, Zhang S, Wang J, Wang J, Yang H, Yang G, Wang J, Paterson A.H, Xia Q, Ji Q, Xiangc Z (2013) Draft genome sequence of the mulberry tree *Morus notabilis*. Natur Commun. 4. DOI:10.1038/ncomms3445.

[27]. Sato S, Nakamura Y, Kaneko T *et al*., Tabata S (2008) Genome structure of the legume, *Lotus japonicus*. DNA Res. 15(4): 227-239. DOI:10.1093/dnares/dsn008.

[28]. Schmutz J, Cannon SB, Schlueter J *et al*., Jackson SA (2010) Genome sequence of the palaeopolyploid soybean. Nature. 463: 178-183. DOI:10.1038/nature08670.

[29]. Varshney RK, Chen W, Li Y *et al*. (2012) Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. Natur Biotechnol. 30: 83-89. DOI:10.1038/nbt.2022

[30]. Young ND, Debelle F, Oldroyd GE *et al*. (2011) The *Medicago* genome provides insight into the evolution of rhizobial symbioses. Nature. 480: 520-524. DOI:10.1038/nature10625

[31]. Price AL, Jones NC, Pevzner PA (2005) De novo identification of repeat families in large genomes. Bioinformatics. 21: 351-358. DOI:10.1093/bioinformatics/bti1018

[32]. Smit AFA, Hubley R, Green P. Repeat Masker Open-4.0. 2013-2015. Available from: http://www.repeatmasker.org

Appendix 1.1

List of primer pairs used to amplify the SSR containing regions, along with information on the individual SSR motifs and LG to which each SRR marker locus belongs. The M13 tailed primers are indicated in bold. The dotted lines separate the multilocus PCRs; the plain lines indicate Genescan® runs.

Old ID[1]	New ID[2]	LG	Motif		Primer sequence	Tm (°C)	GC (%)	PCR cycle	Dye
EU03D01	M7.19	7	$(CT)_{18}$	For	ATGTCGGAGCAAAATCGTTC	59.38	45	Rad-multi 56	6-FAM
				Rev	**TTGTAAAACGACGGCCAGT**CATGTTCCCGCTCATGAATA	58.21	45		
EU10H05	M7.20	7	$(CT)_{31}$	For	**TTGTAAAACGACGGCCAGT**GTCATGATGGCGTAAAAG	53.95	44	Rad-multi 56	
				Rev	ACACTCACTCACACTCCGTAA	61.25	47		
EU02C09	M4.12	4	$(CT)_8TT(CT)_5CC$ $(CT)_3TT(CT)_7$	For	GGCATCGGGATAGAAAAACA	58.43	45	Rad-multi 56	VIC
				Rev	**TTGTAAAACGACGGCCAGT**TCAATGCCTCAACAGAAATCC	59.04	42		
EU07G10	M4.10b	4	$(CT)_{21}CATA\ (CA)_5$ $CT(CA)_5$	For	CATCCATTATTGGGCAG	52.33	47	Rad-multi 54	PET
				Rev	**TGTAAAACGACGGCCAGT**CACCAACGAACTCCTTACAAA	58.92	42		
EU0022	M1.1	1	$(GA)_{40}$	For	**TTGTAAAACGACGGCCAGT**CCAACGGATACCAAGGTGTT	60.45	50	Rad-multi 54	NED
				Rev	AACCGCACGGGTTCTATG	60.18	55		
A124	M9.25	9	$(CA)_{11}$	For	GTGTGGGTGTTTGAAGAGC	59.43	52	Rad-multi 54	
				Rev	**TTGTAAAACGACGGCCAGT**TCAAGAACATCAACGCGTAA	58.31	40		
EU03H01	M4.11b	4	$(TG)_5CG(TG)_7$	For	GCCATTCCTTTCAAGAGCAG	59.66	50	Rad-multi 56	6-FAM
				Rev	**TTGTAAAACGACGGCCAGT**AACCCAAAACCGCAACAATA	59.07	40		
EU01H08	M3.7	3	$(CT)_{22}$	For	TTCGAGTCTTGCCTTAATTGTT	58.75	36	Rad-multi 56	
				Rev	**TTGTAAAACGACGGCCAGT**CAGACGACCTTACGGCAACT	62.64	55		
A149	M3.9	3	$(CA)_{12}$	For	**TTGTAAAACGACGGCCAGT**CTGCTATGGACAGTTCCAGT	59.58	50	Rad-multi 56	VIC
				Rev	CAATTCAGTTGTGATAGACGC	57.81	42		
EU11C09	M3.8	3	$(CT)_{16}$	For	AGGAAGCGGTGTCATCTGT	61.40	52	Rad-multi 56	
				Rev	**TTGTAAAACGACGGCCAGT**CGCCCACATATTCATTCTCA	58.21	45		
sw2H09.2	M2.5	2	$(CT)_5CC(CT)_{13}TT$ $(CT)_5$	For	GTGCCGGTCTTCAGGTTACA	62.60	55	Rad-multi 54	NED
				Rev	**TTGTAAAACGACGGCCAGT**CGCCTACCGATTACGATTGA	60.16	50		
sw2A12.2	M9.27	9	$(GA)_{10}TAAA\ (GA)_5$	For	**TTGTAAAACGACGGCCAGT**GCTAAAAGAAGTGCAAGGAGA	58.70	42	Rad-multi 54	
				Rev	TGTTCTTTCAAGTGCCAA	54.94	38		
B42	M2.6	2	$(CT)_{26}$	For	**TTGTAAAACGACGGCCAGT**GGAGCAGGTAGAGTCCCATC	61.61	60	Rad-multi 54	PET
				Rev	CGTTTGAAAATTTATACCAAAATG	54.54	25		
EU07F12	M6.16	6	$(CT)_{12}TT(CT)_{15}TT$ $(CT)_2TT(CT)_4$	For	**TTGTAAAACGACGGCCAGT**TATTGCATTGTTGTTCCTTG	54.90	35	Rad-multi 54	
				Rev	TGTATTTAGAAGAGGGAAATAGATG	56.44	32		

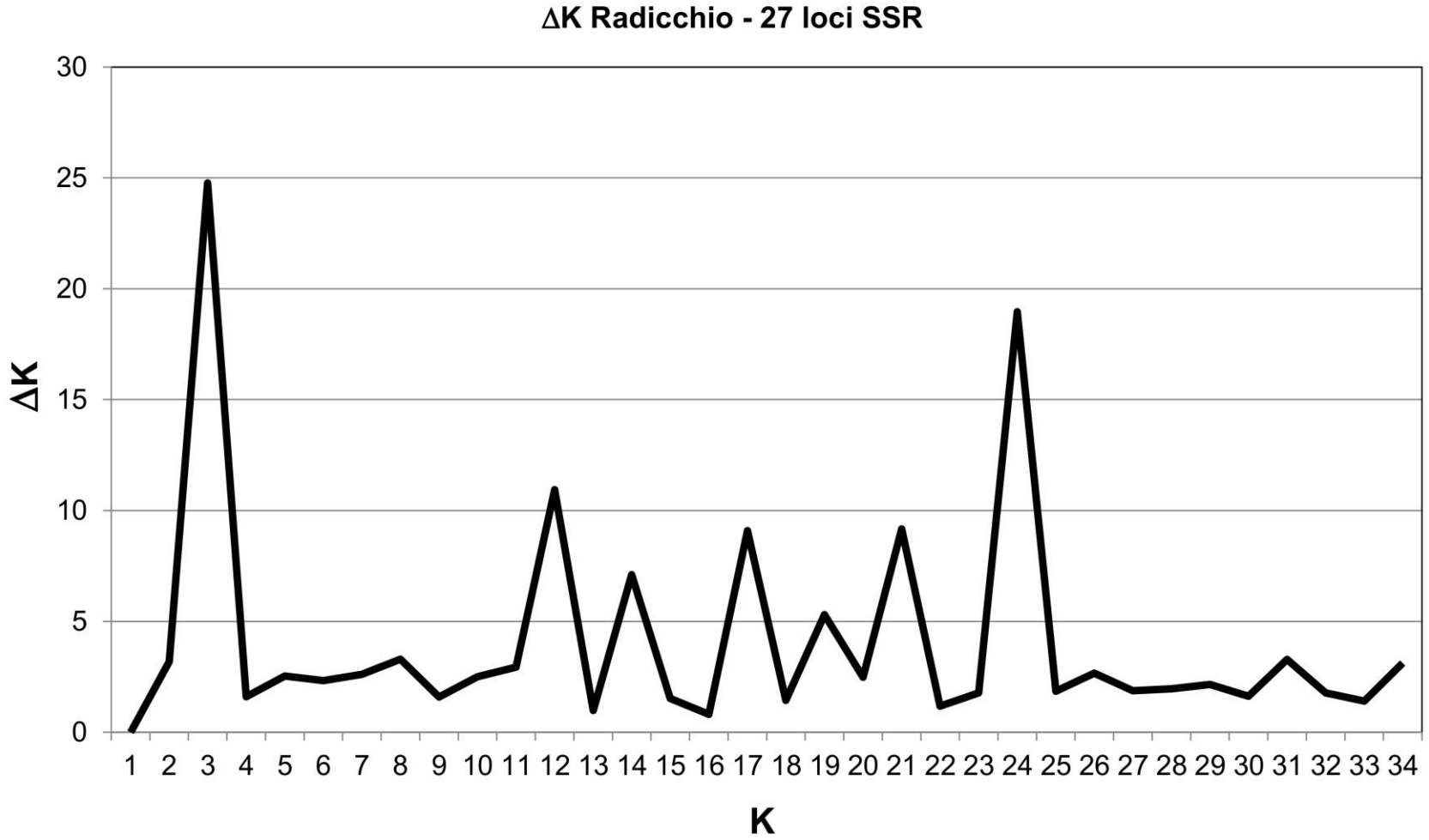
Old ID[1]	New ID[2]	LG	Motif		Primer sequence	Tm (°C)	GC (%)	PCR cycle	Dye
EU07B09	M8.22	8	$(CA)_5AA(CA)_9$	For	TCGTCATCAGAAACAAAGCAA	58.92	38	Rad-multi 56	6-FAM
				Rev	**TTGTAAAACGACGGCCAGT**CAAAGAAGGCACTCTTGTCG	59.60	50		
A94	M8.24	8	$(TC)_{16}(CA)_{13}$	For	**TTGTAAAACGACGGCCAGT**GGTCCGTAGACTGCAGACTTT	62.27	52	Rad-multi 56	
				Rev	CACCGTCCCACTTTTTAGG	58.52	52		
B214	M5.14	5	$(TC)_{11}$	For	**TTGTAAAACGACGGCCAGT**AAAGTCACACATCGCATTTCCT	61.32	40	Rad-multi 56	VIC
				Rev	GTAGCAGCAGCAGCCATCTT	63.38	55		
sw2F09	M9.26	9	$(GATA)_3N_{19}(GA)_9$	For	**TTGTAAAACGACGGCCAGT**CCTACACTCGGCCACCTACT	63.36	60	Rad-multi 56	
				Rev	TCGACGGTATAACAACACCTG	60.28	47		
EU02E02	M7.21	1	$(CT)_{13}$	For	GGACACCGAGCTGGAGAA	61.46	61	Rad-multi 56	
				Rev	**TTGTAAAACGACGGCCAGT**TTCCACTTTCGGGAGTTACC	59.80	50		
EU06C09	M1.3	1	$(CT)_{17}$	For	**TTGTAAAACGACGGCCAGT**TGGAGAAAAATGAAGCAC	53.16	38	Rad-multi 54	PET
				Rev	GAATGAGTGAGAGAATGATAGGG	58.36	43		
B131	M2.4	2	$(GA)_{25}$	For	**TTGTAAAACGACGGCCAGT**GCTCGAAAATCGGCTACAAC	60.09	57	Rad-multi 54	
				Rev	CGAGCCATGTTAGGGTTTGT	60.81	50		
EU08C07	M6.18	6	$(CT)_{16}$	For	CTCAACGAATGCTTTGGACA	59.23	45	Rad-multi 56	6-FAM
				Rev	**TTGTAAAACGACGGCCAGT**CCTCGCGGTAGCTTATTGTT	60.66	50		
EU02A11	M5.13	5	$(CT)_{23}$	For	AGGCATAAAGAGGTGTGG	56.61	50	Rad-multi 56	
				Rev	**TTGTAAAACGACGGCCAGT**TCAAACATGAAAACCGCTC	56.89	42		
EU07C10	M8.23	8	$(CA)_{11}(CT)_9$	For	TGTAGACACACAAAATGCACA	58.84	38	Rad-multi 54	VIC
				Rev	**TTGTAAAACGACGGCCAGT**ACCGGTTGAAAACATGAAAT	56.59	35		
A158	M6.17	6	$(CA)_8(CT)_{18}$	For	**TGTAAAACGACGGCCAGT**CGTGTCCAAACGCAAACATTAT	60.85	40	Rad-multi 54	
				Rev	GCACAATTTTCCTACCACTTATCC	60.42	41		
EU0030	M5.15	5	$(CT)_{11}N_7(CAA)_5$	For	**TTGTAAAACGACGGCCAGT**AGCACGACTCTGCTGTCTTTT	62.75	47	Rad-multi 54	NED
				Rev	CGAGCCATGTTAGGGTTTGT	60.81	50		
EU02D02	M1.2	7	$(CT)_{19}$	For	**TTGTAAAACGACGGCCAGT**CCGGCAGAATTTTTAGGG	56.14	50	Rad-multi 54	
				Rev	CAGGTCATAGGTCCATGTGAAA	60.27	45		

1. Cadalen *et. al.* [10]

2. Ghedina *et. al.*(present study)

Appendix 1.2

Output of the ΔK method showing the two investigated levels of K (*i.e.* K=3 and K=24)



ANDREA GHEDINA

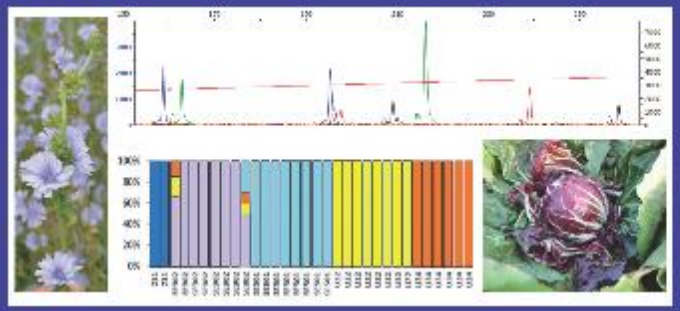
BREEDING F1 HYBRID VARIETIES OF LEAF CHICORY THROUGH MARKER-ASSISTED SELECTION SCHEMES

2016

Università degli Studi di Padova

ANDREA GHEDINA

BREEDING F1 HYBRID VARIETIES OF LEAF CHICORY THROUGH MARKER-ASSISTED SELECTION SCHEMES



Corso di Dottorato in Scienze delle Produzioni Vegetali

2016