Recent insights on Citrus diversity and phylogeny

François Luro, Franck Curk, Yann Froelicher and Patrick Ollitrault
Recent insights on Citrus diversity and phylogeny

François Luro, Franck Curk, Yann Froelicher and Patrick Ollitrault

Citrus trees originated in an extensive area covering Asia (from India to the north of China) and Oceania (Queensland, Australia). The genus Citrus is defined by two different classification systems: Tanaka’s, with 156 species, and Swingle’s, with only 16 species. However, these two systems often contradict each other due to the overall sexual compatibility between the Citrus species and the frequent occurrence of apomixes (due to nucellar polyembryony), which leads many taxonomists to consider interspecific hybrids (vegetatively propagated by apomixes) as new species. The high phenotypic and genetic variability of the citrus taxa reflects a long history of cultivation, in which many mutations and natural hybridizations gave rise to the existing diversity within this mainly facultative apomictic group. Genetic marker studies and complete genome sequence data have recently elucidated the phylogeny of the Citrus genus and especially the origin of edible species.

1. Taxonomy

Citrus species are classified in the Geraniales Order, the Rutaceae Family and the Aurantioideae Subfamily. Aurantioideae has been subdivided into two tribes: 1. Clauseneae with five genera and Citreae with 28 genera including Citrus and related genera, i.e. Fortunella, Poncirus, Eremocitrus, Microcitrus and Clymenia. The tribe Citreae comprises three subtribes: Triphasiinae, Balsamocitrinae and Citrinae; the latter, with 13 genera, has been classified into three groups: 2. group A ‘the primitive citrus fruit trees’ with five genera, Severinia, Pleiospermium, Burkillanthus, Limnocitrus and Hesperethusa; group B ‘near citrus fruit trees’ with only two genera, Citropsis and Atalantia; and group C ‘true citrus fruit trees’ which includes six sexually compatibles genera, Fortunella, Eremocitrus, Poncirus, Clymenia, Microcitrus and Citrus.

The taxonomy of the Citrus genus was, until recently, controversial, complex and sometimes confusing. Two major systems are still widely: the Swingle classification.
considering 16 species (table 1) and the Tanaka (1961) one identifying 156 species. Major horticultural citrus groups such as the orange (C. sinensis (L.) Osb.), mandarin (C. reticulata Blanco), lemon (C. limon (L) Burm.), grapefruit (C. paradisi Macf.), lime (C. aurantifolia (Christm.) Swing.) and pummelo (C. maxima (Burm.) Merr.), are each considered as species in Swingle’s systematics. While Swingle recognized only one species for sweet orange (C. sinensis), Tanaka described 12 species for this citrus horticultural group (table 2). This controversial situation results from the conjunction of a broad morphological diversity, the overall sexual interspecific compatibility within the Citrus genus and between genera, and the partial apomixis of many cultivars. The Citrus apomixis is characterized by the development of somatic (nucellar) embryos in addition to zygotic one. The competition for germination and growth is more favourable for the development of plantlets from nucellar embryos than the promotion of clonal reproduction. Therefore, apomixis fixes and amplifies complex genetic structures by seedling propagation which produces populations of trees with similar phenotypes, consequently considered by taxonomists as new species.5

Table 1 - Taxonomy of Citrus by Swingle (1943).

<table>
<thead>
<tr>
<th>Swingle Systematics (1943)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SECTION</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>Subgenus Citrus</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Subgenus Papeda</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Table 2 - Comparison of sweet orange taxonomy between Swingle and Tanaka systems.

<table>
<thead>
<tr>
<th>Swingle (1943)</th>
<th>Tanaka (1961)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. sinensis</td>
<td>C. sinensis Osbeck</td>
</tr>
<tr>
<td>C. sinensis</td>
<td>C. tankan Tanaka</td>
</tr>
<tr>
<td>C. sinensis</td>
<td>C. temple Hort. ex Tan.</td>
</tr>
<tr>
<td>C. sinensis</td>
<td>C. oblonga Hort. ex Tan.</td>
</tr>
<tr>
<td>C. sinensis</td>
<td>C. funadoko Hort. ex Tan.</td>
</tr>
<tr>
<td>C. sinensis</td>
<td>C. iyo Hort. ex Tan.</td>
</tr>
<tr>
<td>C. sinensis</td>
<td>C. sinograndis Hort. ex Tan.</td>
</tr>
<tr>
<td>C. sinensis</td>
<td>C. luteo-turgida Tanaka</td>
</tr>
<tr>
<td>C. sinensis</td>
<td>C. ujukitsu Hort. ex Tan.</td>
</tr>
<tr>
<td>C. sinensis</td>
<td>C. tamurana Hort. ex Tan.</td>
</tr>
<tr>
<td>C. sinensis</td>
<td>C. aurea Hort. ex Tan.</td>
</tr>
<tr>
<td>C. sinensis</td>
<td>C. shukokan Hort. ex Tan.</td>
</tr>
</tbody>
</table>

Citrus taxonomy is evolving thanks to new information from genetic studies on their phylogeny and diversity. Mabberley has proposed a new classification of edible citrus which recognizes three species and four hybrid groups. However, recent genetic studies shown that even these three classifications are not totally in accordance with the phylogenic history of the citrus.

### 2. Geographical origins

The centres of origin for citrus and its relatives are in southern and eastern Asia, and Australia. Swingle recognized six species; two which are native to Papua New Guinea – Microcitrus. M. papuana and M. warburgiana – and four which are native to Australia. The Australian species of Microcitrus has recently proved to be economic successful due to a fruit called finger lime, commonly known as the caviar lemon. Eremocitrus is a monospecific genus (E. glauca) native to the New South Wales and Queensland deserts (Australia). Clymenia is also a monospecific genus (C. polyandra) and its place of origin is Papua New Guinea. Poncirus is a unique citrus genus distinguished from others by its deciduous leaves; originally from northern China, this citrus tree is the most tolerant to freezing temperatures (resisting up to -20°C).
It was considered a monospecific genus (*P. trifoliata*) for a long time, until two genetic groups were described. In 1984, Ding *et al.* recognized a new species of *Poncirus* (*P. polyandra*). Because of its tolerance to low temperatures, immunity to the *Citrus tristeza* virus and resistance to *Phytophthora* spp., the *Poncirus* genus is directly used or cross combined with other *Citrus* species to produce rootstocks for citrus cultivation. *Fortunella* spp. produce kumquat fruit and depending on the taxonomy, between two and four species are recognized. This genera originated from north-eastern China, making it one of the most cold-tolerant edible citrus trees. The *Citrus* species originated from a large area in south-east Asia. Tolkowsky*¹⁰* considered that the mountainous regions of southern China and north-east India as being their centre of origin. Gmitter and Hu,*¹⁰* however, were more specific and specified the Yunnan province – due to its wide diversity of citrus – as the major centre of origin for the citrus. Tanaka*¹¹* proposed a theoretical dividing line running from the north-western border of India, above Burma, to the Yunnan province of China, and then to south of the island of Hainan (fig. 1). Several citrus species such as citrons (*C. medica*), lemons (*C. limon*), limes (*C. aurantifolia*), pummelos (*C. maxima*) and the sour and sweet oranges (*C. aurantium* and *C. sinensis*) presumably originated south of this line, while mandarins (*C. reticulata*) and others originated north of it. Citrons are indigenous to north-east India, and pummelos to the Malay and East Indian Archipelago.*¹²* The *Papeda* group includes citrus from different geographical origins; *Citrus micrantha* could be native to the southern islands of the Philippines, *C. latipes* to north-east India, *C. macroptera* near to New Caledonia, *C. celebica* to the Indonesian islands, and *C. hystrix*, of an uncertain origin, could be from the Philippines.*¹³*

Fig. 1 - Phylogenetic origins of major secondary Citrus species with the maternal and paternal ancestors (dotted lines are hypothetical cross).
3. Phylogeny of edible *Citrus* species

Despite the difficulties in establishing a consensu al classification of edible *Citrus*, most authors now agree on the origin of cultivated forms. The use of molecular markers such as isoenzymes, RFLP, RAPD, SCAR, AFLP, SSRs, SNP, a mix of Indels/SSR/SNP and genome sequencing have contributed to identifying four basic taxa – *C. maxima* (pummelos), *C. medica* (citrons), *C. reticulata* (mandarins) and *C. micrantha* (a wild Papada species) – as the origin of all cultivated *Citrus*, and in deciphering the genetic origin of the major *Citrus* secondary species. In addition to the nuclear genome investigation, the maternal phylogeny of each cultivated form has been elucidated using the Indel, SSR or SNP markers of their chloroplastic and mitochondrial genomes.

While most modern varieties of pummelos and citrons appear to be pure *C. maxima* and *C. medica*, respectively, recent genomic and molecular marker studies have revealed that almost all modern mandarins are not pure *C. reticulata* but are introgressed by *C. maxima* genome fragments.

A general scheme of phylogenetic relationships between the major *Citrus* species is presented in figure 2. The *C. aurantium* (sour orange) is a direct hybrid between *C. maxima* and *C. reticulata*, where pummelo is the maternal parent. *C. sinensis* (sweet orange) is closer than the sour orange to *C. reticulata* but displays homozygous introgressed fragments of the *C. maxima* nuclear genome; therefore, it cannot be a direct hybrid or a backcross between the ancestral taxa but is probably a second or third generation product. It could be derived from a cross between (*C. maxima × C. reticulata) × C. maxima* as an egg donor and *C. reticulata* as a pollinator, with some introgression with *C. maxima*. *C. paradisi* Macf. (grapefruit) was native of Barbados and introduced to the USA at the beginning of 19th century. It is close to *C. maxima*, but displays alleles from the *C. reticulata* gene pool that are also shared with *C. sinensis*. This could be the result of hybridization between *C. maxima* and *C. sinensis*, with the pummelo as the maternal parent. *C. clementina* (clementine) is a chance seedling hybrid discovered by the Father Clément (V. Rhodier, 1829-1904) at the end of the 19th century in Messerghin (Algeria), close to Oran, in the orchard of an orphanage. This hybrid originated from the fertilization of an ovule of *C. deliciosa* (mandarin) with the pollen of a *C. sinensis* (sweet orange). Tangors and tangelos are horticultural names given to the suspected or controlled hybrids of mandarins (‘tang’ coming from ‘tangerine’ – the name given to mandarins coming from Tangier, Morocco) and sweet oranges, and mandarins and grapefruits, respectively. Their genomes, therefore, are also admixtures of *C. reticulata* and *C. maxima*. 
Recently published work has investigated the diversity and origin of lime and lemon groups by using 123 markers, including 73 SNP markers with specific alleles from the four ancestral species. These diagnostic markers were developed from genomic sequences from across the entire genome provided to identify the origin of different lemon and lime genotypes by calculating the allelic proportion of the four ancestral species (fig. 3). *C. medica* appears to be the male parent of almost all limes and lemons.

*C. limon* (lemon) results from the direct hybridization between *C. aurantium* and *C. medica*, as previously proposed. *C. limetta* (Marrakech limonette) has a similar origin while *C. limettioïdes* (Palestine sweet limes) and *C. meyeri* (Meyer lemon) also display molecular
patterns compatible with a \([C. \text{maxima}/C. \text{reticulata} \text{ add mixture parent}] \times C. \text{medica} \) origin, but with an undetermined female parent.

12 The Mexican lime (\(C. \text{aurantifolia}\)) can be considered as a direct hybrid between \(C. \text{medica}\) and \(C. \text{micrantha}\).\(^3\) Similarly, the model \(C. \text{micrantha} \times C. \text{medica}\) is also applicable for \(C. \text{macrophylla}, C. \text{aurantia}\) and \(C. \text{excelsa}\) though from independent reticulation events. For the New Caledonian and Kaghzi limes, an \(F2 (C. \text{micrantha} \times C. \text{medica}) \times (C. \text{micrantha} \times C. \text{medica})\) origin was proposed.

13 The seedless limes Tahiti, Bearss or IAC (\(C. \text{latifolia}\)) are triploid hybrids resulting from the hybridization between the diploid pollen of \(C. \text{aurantifolia}\) and a haploid ovule of \(C. \text{limon}\). A second group of triploid seedy limes (Tanepao, Coppenrath, Ambilobe and Mohtasseb limes) and the Madagascar lemon had a different phylogenetic origin, probably as the result of a \((C. \text{micrantha} \times C. \text{medica}) \times C. \text{medica}\) hybridization with a diploid gamete from the \(C. \text{micrantha} \times C. \text{medica}\) parent.

14 The names lime and lemon are also attributed to other acidic citrus forms originating from different parental crosses such as the Volkaner lemon, Rough lemon and the Rangpur lime, which initiate from crosses between the mandarin, as maternal parent, and the citron, as pollinator. \(C. \text{bergamia}\) (bergamot) originated in Spain or in the south of Italy around three or four centuries ago, following the fertilization of a sour orange by lemon pollen (\(C. \text{aurantium} \times C. \text{limon}\)).

4. Diversification

15 The phenotypic diversity of the citrus is particularly high, especially in the Asian species, as revealed by molecular markers, chromosomal banding patterns and phenotypic characters – such as fruit pomology and the chemical variability of peel and leaf oils – as well as their tolerance to biotic and abiotic stresses. This is largely due to the evolutionary history of this gene pool and its diversification mechanisms, sometimes specific to each taxonomic group. The diversity studies of morphological, primary and secondary metabolites polymorphisms suggest that a major part of the phenotypical diversity of the edible \(Citrus\) is supported by the ancestral taxa of the cultivated \(Citrus\).\(^3\)

16 The allopatric evolution (geographic isolation) as presented in the 'Geographical origins' section, allowed the ancestral species to diversify by acquiring the specific characteristics of each species, probably conditioned by interaction with the environment of each diversification area. For example, apomixis is only present among taxa whose origin lies north of the Tanaka's line (fig. 1), and only in these taxonomic groups; the skin and pulp are orange coloured due to the synthesis of xanthophyll carotenoids.\(^3\) The flowering period is different between the botanical Chinese genera: in the Mediterranean area \(Poncirus\) bloom in late winter, \(Fortunella\) in the heart of summer and \(Citrus\) usually in the middle of spring. Some reproductive characteristics are also different: pummelos (\(C. \text{maxima}\)) share a strict gametophytic self-incompatibility which imposes cross-fertilization in reproduction, while inbreeding seems to be the preferred reproduction mode of citrons (\(C. \text{medica}\)) which results in the increase of homozygosity.\(^6\) Genome size estimated by flow cytometry is also variable depending on the species (mandarins registered the lowest score, while citrons registered the highest – a 20% increase on the mandarin result).\(^7\) These genome size variations also support the hybrid origins of secondary species, as presented in the previous section.
If sexual reproduction seems to be the main mechanism of diversification within the ancestral species, it is, in contrast, almost absent in the diversification of apomictic secondary species. Nevertheless, the phenotypic diversity of secondary species is also quite important, and is probably the result of somatic mutations events such as SNPs, chromosomal translocations, insertions/deletions, mobility of transposable elements, variation of methylation patterns or changes in the level of ploidy. Butelli et al. demonstrated that the synthesis of anthocyanins – which provides the blood colour to the pulp of some orange varieties (fig. 4) – is related to the insertion of a transposable element in the promoter region of a gene encoding a transcription factor (Ruby gene). The lemon var. *Luninciana* (*C. lemon*) – a very large olive-shaped lemon – differs from Eureka-type lemon varieties by a large deletion located in chromosome 9. The large majority of mutations affecting the phenotype of citrus varieties are of natural origin. However, some crop varieties were obtained by artificial induced mutagenesis (irradiation), which usually made them sterile and produce seedless fruit. This is the case of the Star Ruby grapefruit, which is the product of irradiated Hudson seeds.

Fig. 4 - Phenotypes of sweet oranges varying in fruit seediness and pulp colour (from left to right the half fruits correspond to Parson Brown, Washington Navel, Cara Cara Navel and Moro varieties).

The Giant Key lime is a tetraploid form of the Mexican lime, created by a chromosome doubling in a somatic embryo. The ploidy variation could also affect gametes, ovules or pollen, coming from meiosis dysfunction producing diplogametes, when fertilized by a normal gamete generate triploid offspring. Using SNP diagnostic molecular markers, Curk et al. demonstrated that the genesis of triploid limes were related to the diplogamy in the Mexican lime. Few genomic origins of phenotypic variation have been elucidated, but the phenotypic diversity observed in the secondary species suggests that non-sexual modifications are also relevant diversification mechanisms. The development of new and cheaper genome sequencing methods could provide information which reveals genomic variations helpful to studying their effect on phenotypic diversity.
Conclusion

The broad phenotypic diversity observed in the citrus is most likely a consequence of its large area of diversification combined with a geographically segmented evolution which limited gene flow through populations. During this allopatric evolutionary phase, each population acquired specific characters largely adapted to each environment but without loss of the interfertility capacity between populations. For instance, *Poncirus* originated in the north of China, adapted to freezing temperatures (e.g. bud dormancy, deciduous leaves, early blossom period) yet is sexually compatible with other citrus genera native to sub-tropical or tropical areas. This phase of evolution can be described as that of an incomplete speciation which went on to generate the basic citrus taxa. Later, after extension of the growth area when populations grew in common regions, inter-taxa hybridizations occurred which led to enlargement of the variation of phenotypical traits panel. The characteristics and multiplicity of the phenotypes generally fixed by apomixis in secondary species probably influenced taxonomists to define numerous species.

BIBLIOGRAPHY


Garcia-Lor et al. 2015: A. Garcia-Lor, F. Luro, P. Ollitrault, L. Navarro, Genetic diversity and population-structure analysis of mandarin germplasm by nuclear, chloroplastic and mitochondrial markers, *Tree Genetics and Genomes*, 11, p. 123.


Oueslati et al. 2016: A. Oueslati, F. Ollitrault, G. Baraket, A. Salhi-Hannachi, L. Navarro, P. Ollitrault, Towards a molecular taxonomic key of the Aurantioideae subfamily using chloroplastic SNP diagnostic markers of the main clades genotyped by competitive allele-specific PCR, BMC Genet., 17, p. 188.


NOTES

7. Swingle 1943.
8. Swingle 1943.
23. Wu et al. 2014; Curk et al. 2015; Garcia-Lor et al. 2015.
28. Garcia-Lor et al., 2013; Curk et al., 2015.
29. Luro et al., 2013.
35. Fanciullino et al. 2006a.
42. Ollitrault et al. 2008.
44. Tanaka 1961.

AUTHORS

FRANÇOIS LURO
UMR AGAP INRA-CIRAD Corse, Équipe SEAPAG, station INRA 20230 San Giuliano, France; francois.luro@inra.fr

FRANCK CURK
UMR AGAP INRA-CIRAD Corse, 20230 San Giuliano, France

YANN FROELICHER
UMR AGAP INRA-CIRAD Corse, 20230 San Giuliano, France

PATRICK OLLITRAULT
UMR AGAP INRA-CIRAD, Station de Roujol, 97170, Petit-Bourg, Guadeloupe, France