




OPEN

## Genetic relationships and introgression events between wild and cultivated grapevines (*Vitis vinifera* L.): focus on Italian Lambruscos

A. Schneider<sup>1,2</sup>, P. Ruffa<sup>3</sup>, G. Tumino<sup>4</sup>, M. Fontana<sup>5</sup>, P. Boccacci<sup>1</sup> & S. Raimondi<sup>1,2</sup>


Research efforts on genomic structure and ecology of wild populations of *Vitis vinifera* L. offer insights on grape domestication processes and on the assortment evolution of the cultivated forms. Attention is also paid to the origin of traditional, long-cultivated varieties, often producing renowned and valuable wines. The genetic relationships between 283 *Vitis vinifera* cultivated varieties (subsp. *sativa*) and 65 individuals from 9 populations of the *sylvestris* subspecies mainly from northern Italy were explored by means of molecular markers (27 nuclear and 4 chloroplastic microsatellites). Several episodes of contamination of the wild germplasm by the pollen of specific grape cultivars were detected, implying concern for maintaining the purity of the wild form. At the same time, events of introgression from the wild subspecies resulted playing a crucial role in the emergence of several cultivated varieties with a clear admixed genome ancestry *sativa-sylvestris*. These included Lambruscos originated from the flat areas crossed by the Po and Adige rivers in northern Italy, while other cultivars still called Lambrusco but typical of hilly areas did not show the same admixed genome. Historical and ecological evidences suggesting an adaptative recent post-domestication process in the origin of several Italian Lambruscos are discussed.

Grapevine is the fruit crop of major economic value, covering a surface of around 7.33 Mha worldwide with a production of 75 Mt of grapes mainly used for wine and fresh consumption (statistics 2020 by OIV, <https://www.oiv.int/it/what-we-do/global-report?oiv>).

The Eurasian species *Vitis vinifera* L. is by far the most cultivated. Besides the domesticated subspecies *vinifera* (or *sativa*), its wild relative subsp. *sylvestris* still exists in the temperate areas of Europe, northern Africa, western Asia<sup>1</sup>. Wild *vinifera* grapevines differ from the domesticated vines for three major traits: plant sex (they have dioecious plants while the cultivated are generally hermaphrodite, seldom female), size of bunches and berries (smaller in the wild forms), shape of the pips (more rounded and with a shorter beak in *sylvestris*)<sup>2</sup>. There are no breeding barriers between the two subspecies, so that gene flow can occur in both directions in any overlapping distribution areas.

Archaeological and historical records suggest grapevine domestication took place from the wild progenitor around 8000 years ago at the dawn of the Neolithic, in an area between eastern Anatolia and northern Mesopotamia<sup>3</sup>. Recent genomic evidences based on divergence between wild *sylvestris* and cultivated samples predate to 22,000 years ago the beginning of the domestication process<sup>4</sup>. According to Zhou and collaborators<sup>5</sup> there was a long period of low-intensity management by humans prior purposeful cultivation in the Transcaucasia region. A recent study describing the phenomenon as a whole<sup>6</sup> proposes grapevine domestication took place about 11,000 years ago in two different sites for table and wine grapes, in Western Asia and Caucasus respectively.

In addition to the relocation from a natural to a cultivated habitat close to permanent habitation<sup>7</sup>, domestication included the selection of elite variants, their vegetative propagation by cuttings and further improvement

<sup>1</sup>Institute for Sustainable Plant Protection, National Research Council of Italy (IPSP-CNR), Strada delle Cacce 73, 10135 Turin, Italy. <sup>2</sup>Giovanni Dalmasso Foundation, Largo Braccini 2, 10095 Grugliasco, Turin, Italy. <sup>3</sup>Department of Agricultural, Forest and Food Sciences, University of Turin (DISAFA-UNITO), L. Braccini 2, 10095 Grugliasco, Turin, Italy. <sup>4</sup>Plant Breeding, Wageningen University and Research (WUR), P.O. Box 9101, 6700 HB Wageningen, The Netherlands. <sup>5</sup>Bagnacavallo, Italy  email: [anna.schneider@ipsp.cnr.it](mailto:anna.schneider@ipsp.cnr.it)

by the use of recombinants from natural inter-crosses. One of the key changes in this progression was the breakdown of dioecy, together with the increasing in berry number and size, sugar content and the emergence of differently coloured grapes.

With the spreading of wine culture, domesticated vines expanded westwards for several millennia from the prime (or primes) Western Asian centre (centres). Genetic studies had proved the contribution of the local *sylvestris* germplasm to the development of the cultivated European gene pool by comparing samples from the two *vinifera* subspecies<sup>8–12</sup>. Moreover, the identity between chlorotypes (of maternal inheritance in grapevine) from wild vines and from many local cultivars from the Iberian Peninsula, brought evidence of a secondary domestication event, or at least the introgression from local *sylvestris* into the cultivated populations<sup>13,14</sup>. Other additional domestication centres have been proposed according to a multi-origin hypothesis, with independent lineages developing from wild progenitors in different European areas, as reported in the review by Grassi and De Lorenzis<sup>15</sup>.

On the other hand, events of exchange in the opposite direction, i.e. gene flow from the cultivated compartment towards the wild populations, were also documented<sup>14,16</sup>. These facts, which threaten the survival of pure wild forms, must be promptly recognized. At the same time, also feral accessions, i.e. plants escaped from cultivation, must be clearly identified<sup>17</sup>.

Most of the traditional wine grape varieties currently cultivated arose from natural hybridizations followed by clonal propagation often going on for hundreds of years<sup>2</sup>. Generally, only few main genitors emerged, so that the cultivated pool results often structured in regional groups of related varieties or in larger clusters geographically distributed<sup>18,18–25</sup>. Because of the wide use of propagation by cuttings, only few generations of sexual reproduction occurred in the development of historical grape cultivars<sup>26</sup>. This was also proved by paleo-genomics studies, showing several varieties still grown today first-degree related to genotypes (likely extinct) recovered in Roman sites dating back to I-II century CE<sup>27</sup>.

Populations or isolated individuals of wild *vinifera*, representing what nowadays remains of a previous wide diffusion, have been inventoried and often described in most of the temperate regions from the Iberian Peninsula to northern Iran<sup>28–32</sup>. The observed habitats are alluvial or riparian forests along major rivers and their tributaries, more rarely colluvial areas. The interest towards the wild subspecies, today strongly threatened of extinction, lies in the effort: (a) to preserve the widest possible range of genetic diversity for breeding purpose, (b) to better understand the domestication process and the genetic basis of the major selected traits, and c) to outline the origin of some elite cultivars and their history.

Likely due to the mild climate, a great number of *sylvestris* populations were recorded also across the Italian peninsula and in the major islands<sup>33–35</sup>, indicating Italy as one of the possible refuge areas during the glaciations, and therefore postglacial colonization source. In Italy, in fact, there are abundant and ancient literary references to the wild grapevine, defined as *labrusca* or *lambrusca*. According to Sereni<sup>36</sup>, the oldest literary attestation of *labrusca* is attributed to the Latin author Virgil (70–19 BC) born in Mantua, northern Italy. Other references to wild vines followed, including the one by De Crescentiis<sup>37</sup> from Bologna (in the Emilia-Romagna region) “there are wild vines, called *labrusques*, which grow freely on trees”. The term *lambrusques* to indicate plants of the *sylvestris* subspecies also remains in French<sup>38</sup>, due to the common Provençal/Paleo-Ligurian linguistic matrix<sup>36</sup> geographically referable to southern France and north-western and central Italy. Indeed, it does not seem a coincidence that several grape cultivars called Lambrusco, typical of an area south of the river Po in north-central Italy, clustered in a group intermediate between the two subspecies *sylvestris* and *vinifera*<sup>39</sup> or were considered primitive from an historical point of view, as a transition stage between wild grapes and more evolved cultivated varieties<sup>40</sup>.

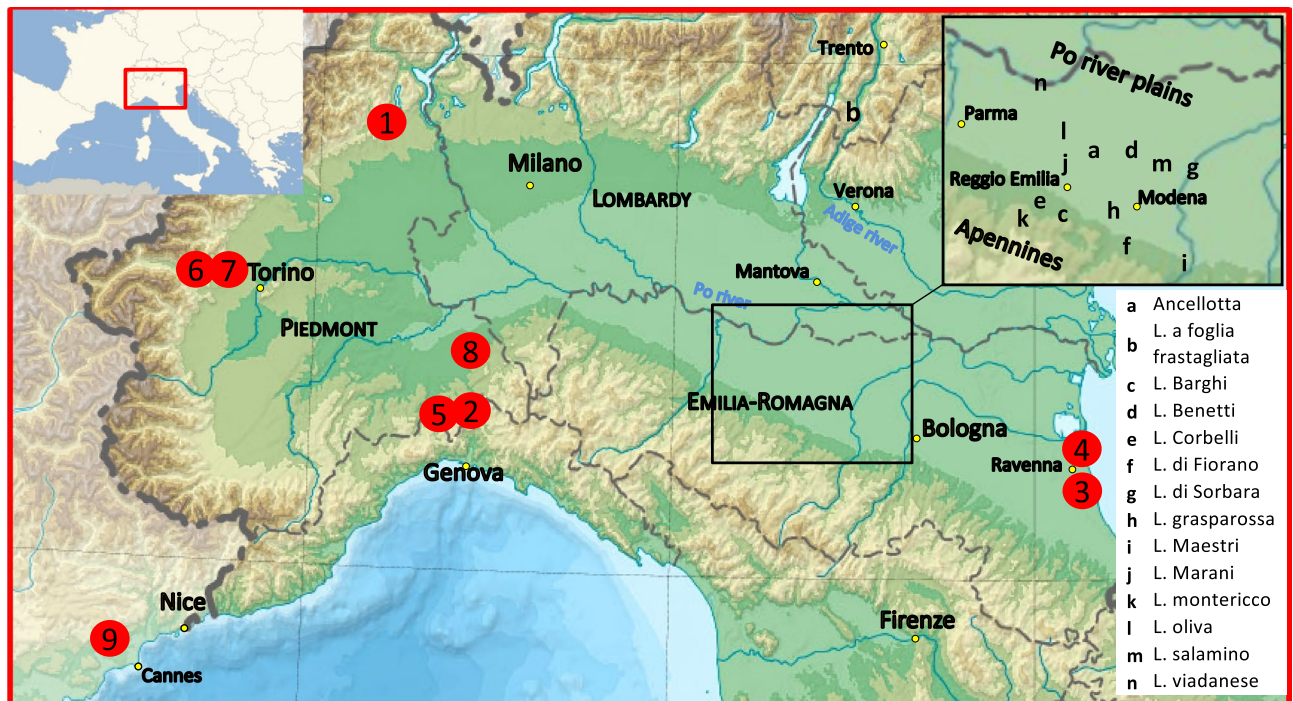
The study we present here explores the genetic relationships among indigenous *sylvestris* populations and local traditional grape cultivars, including 18 named Lambrusco, whose origin from on-site domestication has been proposed but not deeply investigated yet. By means of molecular genotyping (27 nSSR loci well distributed into the genome and 4 cpSSR loci) we analysed 283 cultivated varieties mainly from northern Italy and 65 *sylvestris* individuals from 9 populations located in northern Italy and southern France. Contamination from the domesticated *vinifera* into the subspecies *sylvestris* and the contribution of the latter to the birth of several today's historical cultivars are discussed. An hypothesis about the origin of several Lambrusco varieties denying their ancient ancestry is presented.

## Results

From the 348 accessions examined in this study (283 belonging to cultivated and 65 to wild grapevines, Supplementary Table S1), 348 unique nSSR profiles were obtained (Supplementary Table S2) and chlorotype assigned for 337 accessions (Supplementary Table S1). Of the 18 cultivars named Lambrusco or similarly, most are traditionally cultivated in an area between the towns of Parma, Mantova and Modena, mainly in the region of Emilia-Romagna (Fig. 1). The accessions belonging to the subspecies *sylvestris* were distributed in 9 populations (Fig. 1 and Supplementary Table S3). As to plant sex, some vines showed flowers not clearly classifiable as female or hermaphrodite, bearing both gynoecium and partially, not fully developed stamens (Supplementary Table S1).

## Chlorotype analysis

The most frequent chlorotype in the cultivated section was D, covering 76% of the examined variety profiles (Supplementary Figure S1). If we exclude from the dataset the varieties clearly of non-Italian origin (like Cabernets, Merlot, Pinot, etc.), chlorotype D reach 82% of the domesticated compartment. Chlorotype A is also well recurrent in the Italian cultivated germplasm (16%), while C is uncommon (2%) and B very rare in the set of the examined Italian varieties. As for the 65 wild plants, they are divided between chlorotype A, by far the most frequent (77%), and D for the remainder 23% (Supplementary Figure S1).



**Figure 1.** Geographical distribution of the wild examined grapevine populations (red circles and number according to Supplementary Table S3) and of the traditional growing area of Lambrusco varieties in the Italian region of Emilia-Romagna and Trentino (black letters). Image modified from Blank topographic map from Wikimedia Commons. © Eric Gaba for Wikimedia Commons. Licensed under CC BY-SA 3.0 license ([https://commons.wikimedia.org/wiki/File:Italy\\_topographic\\_map-blank.svg](https://commons.wikimedia.org/wiki/File:Italy_topographic_map-blank.svg)).

### Genetic structure of cultivated and wild *vinifera* grapevines

The analysis by STRUCTURE of the 348 non-redundant nSSR genetic profiles suggested the most likely genetic groups were 2, as inferred by the value of Delta-K (Supplementary Figure S2), the two groups consisting in the cultivated and in the wild plants respectively (Fig. 2a). In this model, several samples from the two sections resulted admixed. Assuming a threshold value 0.5 of the membership coefficient ( $q$ ), the proportion of *sylvestris* ancestry was higher than the domesticated in as many as 13 cultivated genotypes out of 283, while it was lower in one individual collected in the wild (Supplementary Table S4). The  $q$  values of these samples suggested an intermediate ancestry between the two subspecies. They ranged in fact from 0.363 to 0.493 in the group of domesticated for the 13 cultivated genotypes and corresponded to 0.494 in the group of *sylvestris* for the wild plant w-Randolera 4 respectively (Supplementary Table S4).

The 13 varieties genetically closer to the wild populations, ordered by their  $q$  coefficient values in the group of the domesticated were: Lambrusco salamino, Ancellotta, Lambrusco a foglia frastagliata (Enantio N.), Lambrusco Marani, Becuét (Persan), Lambrusco di Corbelli, Chenin, Lambrusco Benetti, Petite Arvine, Lambrusco Maestri, Uvetta di Carpignano, Pinot, Lambrusco viadanese (Fig. 2b). Among other varieties showing a  $q$  value in the cultivated group  $< 0.6$ , thus with admixed ancestry *sativa-sylvestris*, Savagnin blanc (syn. Traminer), Lambrusco oliva, Sauvignon blanc, Lambrusco di Sorbara, Marzemino had an ancestry coefficient  $< 0.52$  (Fig. 2b and Supplementary Table S4).

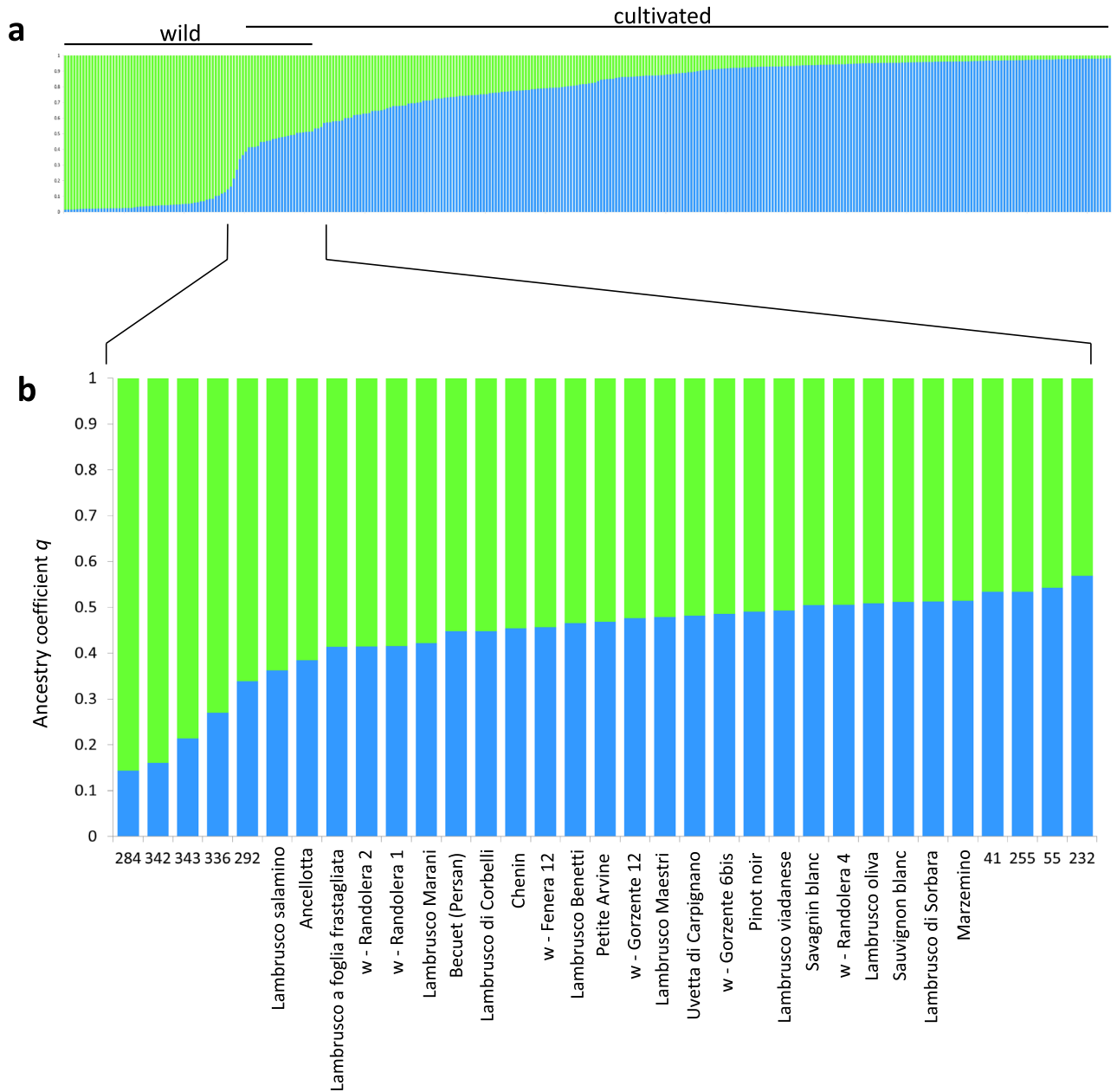
On the side of the grapevines collected into the wild, besides w-Randolera 4 ( $q < 0.5$  in the wild group), the following 5 plants had a  $q$  coefficient  $< 0.6$ : w-Gorzente 6 bis, w-Gorzente 12, w-Fenera 12, w-Randolera 1 and w-Randolera 2 (Fig. 2b and Supplementary Table S4).

PCA outputs from the analysis of 27 nSSR markers indicated a clear separation of wild plants from cultivars along the axis of the major component explaining 17% of the total variance (Fig. 3a). A pre-calculated principal component analysis generated by balanced ancestral populations (see Materials and Methods for details) produced an output diagram (data not shown) consistent with the projection of all samples showed in Fig. 3a.

Some cultivars including many named Lambrusco resulted genetically close to the *sylvestris* branch together with Ancellotta (a major variety from Emilia-Romagna), while other Lambruscos (or similar), notably Lambrusco grasparossa and Lambrusco di Alessandria, were instead distant. On the other hand, several individuals collected into the wild, like w-Gorzente 12 and w-Randolera 4, were allocated among the cultivated plants or on the border between the two groups.

PCA diagram confirmed the ancestry admixture revealed for several samples by STRUCTURE, spatially describing genetic structure of the examined genotypes.

The UWNJ cluster diagram shows a structure similar to the one from PCA with a clear separation of the wild grapevines mostly gathered at one diagram side and the cultivars distributed in several minor groups at the opposite side (Fig. 3b). Here, again, there are some cultivated varieties clustering inside the *sylvestris* group, with Ancellotta and Lambrusco a foglia frastagliata (Enantio N.) the most genetically close to the wild subdivision

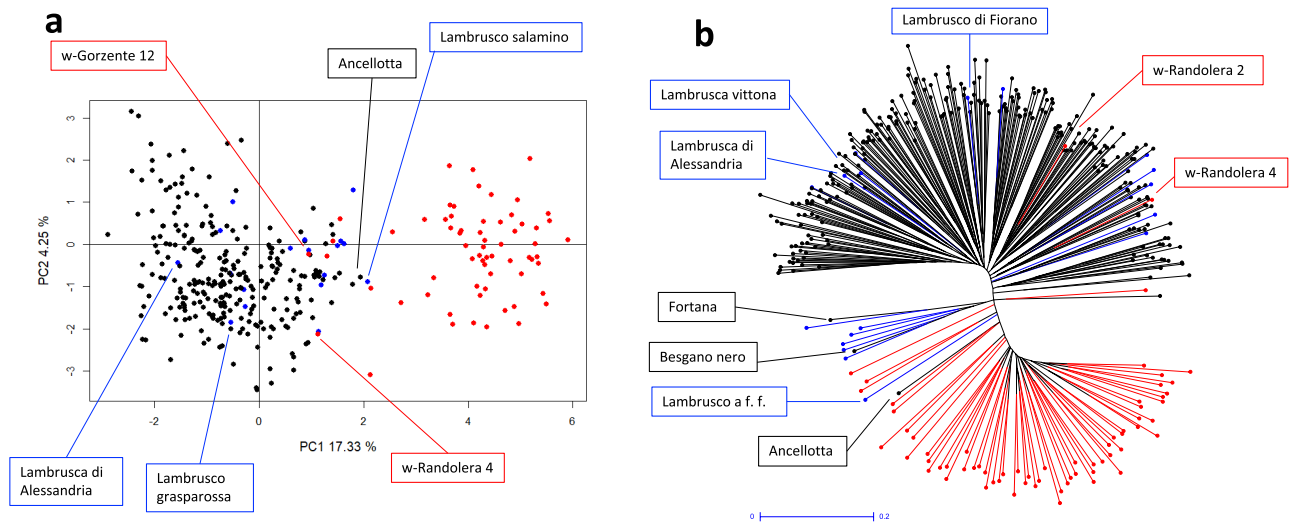


**Figure 2.** Bar plot obtained by STRUCTURE analysing 348 unique genotypes at 27 nSSR markers. Each genotype is represented by a vertical line divided in coloured segments estimating the ancestry membership proportion to STRUCTURE group 1 and 2 (blue and green for the domesticated and the wild subspecies respectively). Two genetic major groups corresponding to the cultivated and the wild grapevines are well noticeable (a). In the enlarged portion of the diagram (b) samples with admixed ancestry referred to in the text are labelled. For the other samples the code number according to Supplementary Table S1 is given.

as also pointed out by STRUCTURE. A cluster that would appear to be transitional between wild and cultivated grapevines includes 8 cultivars among which 6 Lambruscos, with the addition of Fortana and Besgano nero, all traditional from Emilia-Romagna. Parentage analysis (see specific section of this paper) identified these two local varieties as genitors of 6 Lambruscos. This is explained by the fact that the UwNJ analysis favours kinship relationships in the distribution of the samples. Other Lambrusco (like Lambrusco del pellegrino syn. L. di Fiorano), and varieties similarly named (i.e. Lambrusca vittona and Lambrusca di Alessandria) resulted genetically rather distant or even much distant from the wild plants.

On the other hand, two of the wild individuals clustered inside the cultivars, and some were grouped in a sort of intermediate set between the two subspecies (Fig. 3b).





**Figure 3.** Principal Component Analysis (PCA) diagram plot (a) and Neighbour-joining unrooted tree (UwNJ) (b) from 348 non-redundant genotypes analysed with 27 nuclear SSR. In both diagrams, samples from the *sylvestris* populations are in red colour, while black and blue are the cultivated varieties and those named Lambrusco respectively. Only cultivated and wild (w-) samples referred to in the text are labelled.

### Parentage analysis

None of the individuals assumed as wild was found to be a feral accession escaped from cultivation, nor a non-*vinifera* hybrid. This was indicated by both the observations made on plants (leaf and flower morphology) and the analysis of genetic profiles.

Results on the parentage analysis related to the grape cultivars from north-western Italy have been already published by Raimondi et al.<sup>23</sup>. Here we reported the outputs concerning the relationships *sativa-sylvestris* and the cultivated germplasm typical of the region of Emilia-Romagna, from where most of the Lambruscos come as previously mentioned.

Parentage analysis revealed the genotypes collected in the wild showing an admixed ancestry between the two subgroups *vinifera* and *sylvestris* were genetically related to several cultivars. A parent-offspring (PO) link was found between the two individuals w-Randolera 1 and w-Randolera 4 belonging to the wild population Randolera from north-western Piedmont (Fig. 1) and the cultivar Rossara N. (synonyms Schiava N. and Schiava lombarda). The genetic PO relatedness was inferred from high LOD score values and from one shared allele on all the 27 nSSR loci analysed (Table 1 and Supplementary Table S2). Three further nSSR loci analysed were consistent with the mentioned PO (data not shown). The same relationship emerged for plant 2 of the same *sylvestris* population Randolera (w-Randolera 2) and Freisa, an historically widespread variety from Piedmont. Although today, due to the abandonment of grape cultivation in those areas, sporadic familiar vineyards are no less than 2 km distant from the Randolera population site, some decades ago cultivated plots were closer, with a high possibility of contamination of the wild germplasm. With the local name of Uvana we still found in the area isolated, old vines of the cultivar Rossara N. (syn. Schiava), while Freisa is still present mixed up to other varieties in many local longstanding vineyards.

The contamination by domesticated forms were observed for wild individuals from two other *sylvestris* populations. The wild grapevine w-Gorzente 6bis, from southern Piedmont, resulted linked to Dolcetto, one of the major Piedmontese variety (Table 1). Fenera plant 12 was PO related with the cultivar Cagambraga, an ancient, today severely endangered variety recently recovered just in one site in north eastern Piedmont. Although this place is a few tens of km away from the wild grapevine population site, it is certainly conceivable that this cultivar, which has now disappeared, was more widespread in the past.

Ultimately, a contaminant specific domesticated variety was found for all the wild genotypes of admixed ancestry according to STRUCTURE.

Coming to the group of the cultivated varieties whose ancestry according to STRUCTURE was shared with the wild subspecies, only in few cases parentage analysis suggested a possible ancestor. As expected, none of the analysed wild individuals show a genetic profile consistent with a parentage likelihood for these varieties, despite their nearly 50% of ancestry attributable to the *sylvestris* genome. This can be easily explained by the fact that the populations of wild grapevines examined today are nothing but a fragmented remnant of the original wild germplasm likely much more expanded and differentiated at the time of the rising of the cited domesticated varieties.

As to the Lambruscos from Emilia-Romagna, Lambrusco Benetti, Barghi, Marani, salamino and oliva were all suggested as offspring of Besgano nero (Table 1), an ancient grape of the same rural areas, likely being a *sylvestris* individual (or a *sylvestris* descendant) the other parent. A possible descendant from a wild plant inter-crossed with Verdicchio bianco (a variety spread from the Italian Adriatic coast to northern Italy) could be Lambrusco di Corbelli.

For other cultivars resulted genetically close to the wild subspecies, PO relationships were confirmed for Buecuet (syn. Persan), Bian ver (syn. Verdesse), Chenin, Pinot and Sauvignon as likely descendant from Savagnin,

| Grapevine 1 (Gv1)                              | Chlorotype Gv1 | Grapevine 2 (Gv2)                           | Chlorotype Gv2 | Pairs of nSSR loci compared | Mismatching nSSR loci | Pair LOD score | Literature    |
|--|----------------|---|----------------|-----------------------------|-----------------------|----------------|---------------|
| <b>A</b>                                       |                |   |                |                             |                       |                |               |
| Rossara N. (Schiava N.)                        | D              | w-Randolera 1                               | A              | 27                          | 0                     | 19.6           |               |
| Rossara N. (Schiava N.)                        | D              | w-Randolera 4                               | A              | 27                          | 0                     | 18.8           |               |
| Dolcetto N                                     | D              | w-Gorzente 6bis                             | A              | 27                          | 0                     | 17.8           |               |
| Cagambraga                                     | D              | w-Fenera 12                                 | A              | 27                          | 0                     | 17.6           |               |
| Freisa N                                       | D              | w-Randolera 2                               | A              | 27                          | 0                     | 13.7           |               |
| <b>B</b>                                       |                |   |                |                             |                       |                |               |
| Besgano nero                                   | D              | Besgano bianco                              | n.d            | 27                          | 0                     | 18.2           |               |
| Besgano nero                                   | D              | Lambrusco Oliva N                           | A              | 27                          | 0                     | 17.5           |               |
| Besgano nero                                   | D              | Lambrusco salamino N                        | A              | 27                          | 0                     | 17.1           | <sup>19</sup> |
| Besgano nero                                   | D              | Lambrusco Marani N                          | D              | 27                          | 0                     | 16.6           | <sup>19</sup> |
| Besgano nero                                   | D              | Lambrusco Barghi N                          | A              | 27                          | 0                     | 22.4           |               |
| Besgano nero                                   | D              | Lambrusco Benetti N                         | A              | 27                          | 0                     | 14.1           |               |
| Coccalona nera (Orsolina, Rohrtraube blauroth) | A              | Lambrusco del pellegrino N. (L. di Fiorano) | A              | 27                          | 0                     | 19.4           | <sup>23</sup> |
| Coccalona nera (Orsolina, Rohrtraube blauroth) | A              | Parmesana                                   | D              | 27                          | 0                     | 18.8           | <sup>23</sup> |
| Coccalona nera (Orsolina, Rohrtraube blauroth) | A              | Cavazzina                                   | D              | 27                          | 0                     | 13.4           |               |
| Fortana N                                      | D              | Lambrusco Maestri N                         | A              | 27                          | 0                     | 23.9           | <sup>41</sup> |
| Fortana N                                      | D              | Uva toska N                                 | D              | 27                          | 0                     | 12.2           | <sup>41</sup> |
| Rossara N. (Schiava N.)                        | D              | Uvetta (from Carpignano S.)                 | n.d            | 27                          | 0                     | 13.5           |               |
| Uva toska N                                    | D              | Uva toska bianca                            | D              | 24                          | 0                     | 18.8           | <sup>25</sup> |
| Uva toska N                                    | D              | Lambrusco Montericco N                      | D              | 27                          | 0                     | 11.1           | <sup>41</sup> |
| Verdicchio bianco B                            | D              | Lambrusco di Corbelli                       | D              | 27                          | 1*                    | 14.8           |               |
| <b>C</b>                                       |                |   |                |                             |                       |                |               |
| Cabernet franc N                               | D              | Cabernet sauvignon N                        | D              | 27                          | 0                     | 29.7           | <sup>42</sup> |
| Cabernet franc N                               | D              | Merlot N                                    | C              | 27                          | 0                     | 20.9           | <sup>19</sup> |
| Garganega B                                    | D              | Trebbiano toscano B                         | D              | 27                          | 0                     | 24.7           | <sup>43</sup> |
| Gouais blanc (Liseiret B.)                     | C              | Gamay N                                     | C              | 27                          | 0                     | 19.1           | <sup>44</sup> |
| Gouais blanc (Liseiret B.)                     | C              | Riesling B                                  | A              | 27                          | 0                     | 18.3           | <sup>45</sup> |
| Gouais blanc (Liseiret B.)                     | C              | Chardonnay B                                | C              | 27                          | 0                     | 15.0           | <sup>44</sup> |
| Heptakilo                                      | B              | Zibibbo B. (Muscat of Alexandria)           | B              | 27                          | 0                     | 30.3           | <sup>19</sup> |
| Moscato bianco B                               | D              | Zibibbo B. (Muscat of Alexandria)           | B              | 27                          | 0                     | 19.7           | <sup>46</sup> |
| Pinot noir N                                   | A              | Chardonnay B                                | C              | 27                          | 0                     | 24.8           | <sup>44</sup> |
| Pinot noir N                                   | A              | Gamay N                                     | C              | 27                          | 0                     | 15.1           | <sup>44</sup> |
| Sauvignon blanc B                              | D              | Cabernet Sauvignon N                        | D              | 27                          | 0                     | 19.1           | <sup>42</sup> |
| Savagnin blanc B                               | D              | Pinot noir N                                | A              | 27                          | 0                     | 20.8           | <sup>19</sup> |
| Savagnin blanc B                               | D              | Bian ver B. (Verdesse B)                    | D              | 27                          | 0                     | 17.8           | <sup>19</sup> |
| Savagnin blanc B                               | D              | Sauvignon blanc B                           | D              | 27                          | 0                     | 15.9           | <sup>19</sup> |
| Savagnin blanc B                               | D              | Chenin B                                    | D              | 27                          | 0                     | 14.8           | <sup>19</sup> |
| Savagnin blanc B                               | D              | Becuet N. (Persan N)                        | A              | 27                          | 0                     | 14.6           | <sup>19</sup> |

**Table 1.** List of the parent/offspring relationships (PO) resulted from this study between: **(A)** cultivated varieties and wild plants, **(B)** cultivated varieties from the region of Emilia-Romagna, and **(C)** reference PO well documented by literature. Only PO with 0 mismatching loci (except the duo Verdicchio bianco/Lambrusco di Corbelli) and scoring LOD > 11.0 are presented. \*Mismatching is due to a null allele at locus VvMD36.

while for other varieties of high *sylvestris* ancestry, like Ancellotta, Lambrusco a foglia frastagliata and Petite Arvine, no possible relative was suggested.

The lineage of other varieties named Lambrusco (or similarly designated) but distant from the *sylvestris* subspecies, like Lambrusca di Alessandria, Lambrusca pignata and Lambruschetta, all from Piedmont, were already

examined by<sup>23</sup>. Lambrusco Maestri as likely descendant from Fortana N., and Lambrusco Montericco from Uva toscana confirmed previous findings<sup>25,41</sup>. Cocolona nera (syn. Orsolina, Rohrtraube blaurot) was confirmed one of the major main ancestors in northern Italy.

## Discussion

The distribution of chlorotypes detected in this study confirms what has already been observed in the Italian peninsula<sup>13</sup> both for *sylvestris* grapevines (with a predominance of the type A and a considerable presence of D) and for cultivars, where the two chlorotypes are predominant but with an inverse ratio (Supplementary Figure S1).

Péros et al.<sup>47</sup> suggested the species *V. vinifera* arose from the hybridization of two lineages (carrying chlorotypes A and D respectively) expanding from Asia to Europe in separate dispersal events. Types B and C, absent in the Asian species, would develop from A. These findings seem to explain the absence of chlorotype D at the western edge of Europe, and its increasing presence in Europe (together with B and C) from the Italian peninsula moving eastward towards Asia, as well indicated by chlorotype distribution depicted by Arroyo-García et al.<sup>13</sup>.

Unlike other contexts, where it was suggested for some individuals collected in the wild seed dispersal from vineyards<sup>24</sup>, our admixed wild plants shared approximately half of the genome with *sylvestris* subspecies and half with specific varieties. Sangiovese and Trebbiano toscano, leader cultivars in central Italy, were similarly identified as parents of wild accessions in Tuscany<sup>16</sup>.

In the first studies on genetic relationship between the two *vinifera* subspecies, a putative event of secondary domestication has been stated in the Italian island of Sardinia by Grassi et al.<sup>39</sup>. The two involved cultivars Bovale murru and Bovale muristellu (identical to the Spanish varieties Graciano and Parraleta respectively) shared 50% of their alleles at 6 nSSR loci with 3 individuals of a *sylvestris* population. Actually, data from the cited pioneering work indicate more likely an example of wild population contamination by cultivated allochthonous grapevines.

For *sylvestris* populations located in north-western Italy, putative pollen donors were Freisa, Rossara (syn. Schiava or Schiava lombarda) and Cagambraga. The latter two varieties, currently endangered in north western Italy, were recovered in old local vineyards for family use, at the time rather distant from the sites harbouring wild grapevines. Indeed, wild plants can preserve in their descendant generations the evidence of the presence in the past of an intense viticulture now gone. Also, the discover of such pollen donors now rare, was possible thanks to the recovery of the threatened local *sativa* germplasm. As to the admixed wild vines from Gorzente population, the contaminant was likely Dolcetto, one of the major wine grapes of the area. Again, wild plants seem to keep the signs of an impending viticulture for years.

Observations on flower morphology of individuals resulting admixed indicated W-Randolera 4 and w-Gorzente 6 bis showed hermaphrodite flowers, with fully developed stamens and ovary (Supplementary Table S1). In other plants (w-Randolera 1, w-Randolera 2 and w-Fenera 12) a normal ovary with short and straight stamen filaments (instead of revolute as normally in female flowers) were observed. It has not been ascertained whether these flowers were physiologically female or hermaphrodite, although short filaments do not normally go along with anthers developing fertile pollen.

Gene flow from domesticated forms to wild populations had been already reported in different regions and under different ecological conditions<sup>9,11,14,48</sup>. Di Vecchi and collaborators<sup>49</sup> investigated the process by a paternity-based approach, analysing for their parentage 305 seedlings obtained from female wild plants in different locations. The estimated pollen immigration from the cultivated compartment ranged from 4.6% to as much as 26%, indicating the wild germplasm can be severely affected in its evolution with risk of being denatured and in the long run disappearing. In northern Spain, the wild grapevine populations described suffered a dramatic regression in the 20-year period between prospectings, with a significant increase of feral accessions<sup>50</sup>. This is particularly alarming for grapevine genetic diversity preservation. In fact, *sylvestris* forms in compare to the domesticated have been shown to possess genomic regions significantly enriched of functional gene classes related to response to biotic and abiotic stress and thus valuable source of resilience in grapevine breeding programs<sup>51</sup>.

The genetic proximity of several grape cultivars to wild *vinifera* forms, when not explained by the contamination of the latter by the former, indicate events of introgression in the development of local cultivated genetic pool. A genomic approach gave evidence of the presence of an introgression signature in wine varieties across Western Europe and particularly in a subset of Iberian varieties<sup>52</sup>. When the geographic dispersal of a nascent crop to new locations arose, adaptation to local environments is often facilitated by introgression between the crop and locally adapted wild populations<sup>5</sup>. A process of adaptive introgression, although not supported in our study by strict genomic insights, likely occurred in the emergence of several wine grapes examined here.

This seems evident for sought European varieties like Cabernet franc and its descendants (Cabernet sauvignon and Merlot), already found close to the wild genotypes<sup>9,10,16</sup>. As already Bronner<sup>53</sup> had assumed, in our study Savagnin (syn. Traminer) was confirmed to share an ancestry percentage of around 50% with the *sylvestris* subspecies, with an even higher degree of wild ancestry in its descendants Becuét, Chenin, Pinot and still high in Sauvignon.

Petite Arvine, an ancient reputed wine grape of likely Swiss origin<sup>54</sup>, resulted also highly admixed with the *sylvestris* subspecies. The considerable high vigour of its shoots of long internodes and the berries of reduced size with a minimal yield in must, are all characteristics that cannot failed to recall the forms of wild grapevine.

Coming to Lambruscos (or similarly named cultivars), only those cultivated in the region of Emilia-Romagna (and indeed not all of them) showed a high level of wild ancestry, together with another grape traditional of the same area, Ancellotta.

Therefore, Lambrusca di Alessandria, Lambruca vittona, Lambrusca pignata and Lambruschetta from north western Italy recall in their name the Latin term *lambrusche* for some characters reminding the wild grapevines like the high content of pigments and tannins in their fruits.

An origin by an active process of introgression of genomic regions from the wild compartment appear consistent for Ancellotta and for 9 out of 14 Lambruscos from Emilia-Romagna (including Lambrusco a foglia frastagliata more cultivated in Trentino). In this regard, it appears significant that 4 out of the 9 mentioned Lambruscos exhibit a female flower considered to be an ancestral trait in *vinifera* evolution<sup>40</sup>.

Several Lambruscos, as well as Ancellotta, showed a higher ancestral wild grapevine component among a great set of worldwide spread varieties<sup>6</sup>. Similarly using SNP markers, Lambrusco di Sorbara and L. a foglia frastagliata (Enantio N.) were found genetically close to the wild section also by Magris et al.<sup>12</sup>, while in the same work the other examined Lambrusco (L. grasparossa) was not. Accordingly, our analyses showed Lambrusco grasparossa, L. Barghi, L. Montericco and L. del pellegrino (syn. Lambrusco di Fiorano) had an estimated lower wild ancestry. Despite their name, the origin of this second set of Lambruscos does not seem to directly involve the *sylvestris* subspecies.

As to the emergence of Lambrusco varieties, the fermentation of wild grapes is documented in Emilia-Romagna (the Italian Lambrusco's region) in XVII century<sup>55</sup>. The transport of wild *lambrusche* grapes, "black and robust", from Modena and Sorbara to the Ducal cellar to make wine with, is documented in 1693 and 1748<sup>56</sup>.

It is not clear when the term to designate wild plants became the name of cultivated varieties. Historical local chronicles report on a very cold period between 1600 and 1700, leading to the death of many cultivated vines in the region. It is perhaps that seedlings derived from the contribution of local *sylvestris* began to be used for their rusticity and cold resistance to renew vineyards. What is certain is that Lambruscos as cultivated varieties appeared rather recently in the historical grapevine panorama, generally not before the beginning of XIX century (Table 2). Their names refer to the place where the grape was popular or presumably emerged (Sorbara, Montericco, Fiorano, Viadana), the surname of the person or family to whom the introduction or spreading is owed (Barghi, Benetti, Corbelli, Maestri) or a specific grape feature (grasparossa for the red stem, salamino for the compact long and narrow bunch, oliva from olive for the elongated berries). These specifications arose rather recently, when it became important to distinguish different types within the Lambrusco family.

The emergence of Lambrusco as cultivated varieties must also deal with the preference of water flooded areas by the western *sylvestris* vines. The historical area of the Lambruscos related to wild *vinifera* are the flooded lands with emerging groundwater around the rivers Po in Emilia-Romagna and Lombardy, and Adige in Trentino (Fig. 1). Until the recent past these sites were neither reclaimed nor sheltered by river banks. The presumed historical landscape of those areas in the centuries between 1500 and 1700 consisted of cultivated fields separated by rows of trees on which domesticated vines grew, interspersed in the more damped sites with remnants of the original forest<sup>57</sup>.

| Variety name                                  | Current surface (hectares on 2022) | First known historical reference     |  | Synonyms/remarks  |
|---|------------------------------------|--------------------------------------|--|---|
|   |                                    | Where                                | When   |   |
| Ancellotta                                    | 4900                               | Massenzatico (RE)                    | Dalla Fossa, 1810                                      | Lancellotta   |
| Lambrusco a foglia frastagliata (Enantio N.)  | 11                                 | Avio (TN)                            | Comitato Vitivinicolo della Provincia di Trento, 1954  | Enantio   |
| Lambrusco Barghi N. *                         | 4                                  | Albinea (RE)                         | Early 1800s (Rinaldi and Valli, 1992)                  |   |
| Lambrusco Benetti N                           | 3                                  | Carpigiano (MO)                      | Early 1900s (oral communication)                       | Lambrusco Benatti   |
| Lambrusco del pellegrino N. (L. di Fiorano) * | 0.7                                | Fiorano (MO)                         | Vicini, 1752   | Lambrusco di Fiorano, Lambrusco, Lambrusco oliva grosso                 |
| Lambrusco di Corbelli                         | –                                  | Rivalta (RE)                         | Pizzi, 1891  | Lambrusco di Rivalta  |
| Lambrusco di Sorbara N                        | 1300                               | Sorbara (MO)                         | Bertozzi, 1840   | Lambrusco dalla viola, Lambrusca  |
| Lambrusco grasparossa N. *                    | 2000                               | Castelvetro (MO)                     | Aggazzotti, 1867                                       |   |
| Lambrusco Maestri N                           | 800                                | Rubiera (RE)                         | 1854 (reported in Sestini and Fabrini, 1862)           | Dr. Agostino Maestri (from Rubiera) cultivated and spread this variety  |
| Lambrusco Marani N                            | 490                                | Plain area of Reggio Emilia province | Pizzi, 1891  | Lambrusco Barani  |
| Lambrusco Montericco N.*                      | 19                                 | Montericco (RE)                      | Pizzi, 1891  | Selvatica di Montericco, Lambrusco selvatico, Lambrusco di Montericco   |
| Lambrusco oliva N                             | 200                                | Sorbara (MO)                         | Maini, 1851  | Olivina nera, Lambrusco mazzone (RE), Olivone (MN), Groppello, Grepello |
| Lambrusco salamino N                          | 4800                               | Carpigiano (MO)                      | Mid 1700s (Paltrinieri P., reported in Bertozzi, 1840) | Salamina  |
| Lambrusco viadanese N                         | 6                                  | Viadana (MN)                         | Acerbi, 1825   | Lambrusco mantovano, Grappello Ruberti                                  |

**Table 2.** The Lambrusco-named varieties and Ancellotta from Emilia-Romagna and Trentino examined in this study. \*Lambrusco Barghi, L. del pellegrino, L. grasparossa and L. Montericco showed lower *sylvestris* ancestry. MN Mantova province, MO Modena province, RE Reggio-Emilia province, TN Trento province. Literature references of this table are available in Supplementary Table S6.



Associating ancestry coefficients with bioclimatic variables, Magris et al.<sup>12</sup> showed that the genetic component related to western *sylvestris* (wild grapevine from Europe) was significantly and positive associated to seasonal precipitations while the opposite was for the eastern *sylvestris* (wild grapes from the Caucasus regions). These results confirm what reported in pioneering studies on wild *Vitis vinifera* ecology, when it was observed that from the Caucasus towards east the local subspecies, unlike the European *sylvestris*, would prefer arid soils<sup>58</sup>. These observations have found clear confirmation in a recent genomic work on grapevine domestication<sup>6</sup>, where hundreds worldwide samples from the *sylvestris* subspecies clearly clustered in two main groups of European and western Asian origin respectively.

The *sylvestris* subspecies should have introgressed into the new seedlings also genomic regions linked to resistance to bunch rots, whose development is highly facilitated by humid ecological conditions. Among the distinctive regions of the genome between wild and cultivated samples and even within the introgressed tracts, genes connected to the hormonal functionality of plant water regulation and pathogen response have been demonstrated<sup>52</sup>. It is probably no coincidence that most of the Lambruscos exhibit high content in grape anthocyanins and often in tannins. Likely for the same reasons, the French vineyards in the Pyrenean northern slope exposed to Atlantic humid climate harbour native grape varieties with deeply coloured and tannic grapes, such as Tannat and Petit Manseng. These two varieties were both found to be genetically related to the *sylvestris* compartment<sup>10,16</sup>, while Tannat introduced in the humid Uruguayan environment became soon the most popular variety there.

Interestingly, the 4 Lambrusco with less related ancestry to *sylvestris* are varieties mainly developed and still grown in the southern hilly areas of the region, further away from the flooded plain of the Po river (Fig. 1).

As to the Lambrusco's lineage, the cultivar that most contributed to their origin is Besgano nero (Table 1). Besgano nero is an ancient, at least five century-old grape variety (first mentioned by Gallo<sup>59</sup>), with abundant, juicy berries mainly used for fresh consumption. In the past it was often planted together with trees in single rows dividing the cultivated plots along the middle and lower course of the river Po. Its wide spreading in the plain landscape of the region explains its massive role to the origin of many Lambruscos. Besgano's chlorotype (D) indicates it was the pollen donor for 4 out of its 5 Lambrusco descendants (Table 1), which shared their type A with by far the most of the *sylvestris* examined plants. Thus, for the 3 Lambruscos (L. oliva, salamino and Benetti), whose wild ancestry resulted around 50%, the state of seedlings collected by farmers at the edges between cultivated fields and forests (if not into the forests) appears consistent.

In conclusion, the cultivars Lambrusco from Emilia-Romagna of admixed genome *sativa-sylvestris*, should be considered examples of a recent process of adaptive introgression from *Vitis vinifera* subsp. *sylvestris* in the cultivated gene pool. This process has made it possible to gain areas for viticulture to which most of the domesticated varieties were not adapted. A greater insight into the functionality of genomic regions involved will explain the selective processes underlying the origin of these Lambruscos. Not proto- or primitive varieties, but examples of active and recent post-domestication.

## Materials and methods

### Plant material

The accession panel (348 non-redundant genotypes, Supplementary Table S1) comprised: (a) a wider group of traditional cultivars (not issued from breeding) from northern Italy and the Italian north central region of Emilia-Romagna, either widely grown or minor varieties and even neglected or at risk of extinction; 18 varieties named Lambrusco (or similarly named) were included in this group; (b) international cultivars among which some like Savagnin blanc, Pinot noir and Cabernet franc already found genetically related to the *sylvestris* subspecies; (c) important traditional cultivars from the Mediterranean area (e.g. Grenache noir, Carignan noir, Muscat à petits grains blancs, Aglianico, etc.) used as outgroup; (d) 65 *V. vinifera* individuals collected in the wild.

Samples consisting in fresh tissue from young shoots (around 2 g) were collected: (a) in the *ex-situ* collection of Grinzane Cavour (<http://www.ipsp.cnr.it/en/thematics/turin-headquarter-thematics/grinzane-cavour/>) for cultivars from north-western Italy, from the Mediterranean area and international; (b) in commercial vineyards (often old plots) for cultivars from the Emilia-Romagna region, including those named Lambrusco, recovered during a project aimed to the varietal identification of local grape germplasm<sup>60</sup>; (c) in their natural place of discovery for wild grapevines from 9 populations located in northern Italy and southern France (Fig. 1 and Supplementary Table 1). Sample collection was in accordance with the relevant regulations and authorizations. Those from the wild subspecies were collected and identified by authors. Six out of the 9 wild populations were previously inventoried and described<sup>61</sup>. Wild samples from 2 sites in the region of Emilia-Romagna (Pineta di Classe and Pineta San Vitale) were recovered and included in the study because remnants of populations from the same region were the cultivars Lambrusco likely developed. We based our surveys on reports of the presence, albeit sporadic, of *V. vinifera* subsp. *sylvestris* in the Ravenna's historical pine forests<sup>62</sup>.

### Genotyping

DNA extraction was performed using the protocol described by Thomas et al.<sup>63</sup>, starting from young leaves. All the samples were analysed at 27 nSSR loci (Supplementary Table S5). The first 9 SSR markers were those recommended by the international scientific community and include: the 6 loci selected as genetic descriptors for *Vitis* by the International Organization of Vine and Wine (OIV, <https://www.oiv.int/sites/default/files/2022-12/Code%20e%20edition%20Finale.pdf>) and the 3 markers suggested as common reference for grape variety identification within the European *Vitis* Database ([www.eu-vitis.de](http://www.eu-vitis.de)—“Descriptors/file format—Detailed SSR-marker specific information”). The remaining 18 nSSRs were selected based on their polymorphism and coverage of the 19 chromosomes.

Four chloroplast SSR (cpSSR) markers, sufficient to discriminate the 4 chlorotypes commonly found in *V. vinifera*, were also examined (Supplementary Table S5). The Polymerase Chain Reaction (PCR) conditions used for both nSSRs and cpSSRs are described by Ruffa et al.<sup>64</sup>.

### Ampelographic characterization and trueness-to-type

The examined cultivated grapevines were identified by mean of plant morphology, historical descriptions and their genetic profiles from literature and from the following reference databases: (a) The Italian *Vitis* Database, <https://vitisdb.it>; (b) VIVC *Vitis* International Variety Catalogue <https://www.vivc.de>; (c) The Italian Catalogue of grapevine varieties <http://catalogoviti.politicheagricole.it/catalogo.php>.

All the varieties presumably belong to the *vinifera* species, except for Jacquez, a natural interspecific hybrid included in the analysis because recovered in the key region of Emilia-Romagna, where it was mistaken for a local traditional grape.

All the wild sampled vines were carefully observed and described by expert ampelographers, in order to avoid including, at least as far as revealed by morphology and ecology of the plants, feral individuals escaped from cultivation or hybrids from species other than *vinifera*. Flower sex was also recorded.

### Analyses of population structure

Genetic structure was explored based on the nSSR profiles at 27 loci and using a model-based approach, Principal Component Analysis and hierarchical clustering, as detailed below.

The software STRUCTURE v. 2.3.4<sup>65</sup> was used to infer the membership of the 348 non-redundant grapevine genotypes to ancestral groups. A number of hypothetical ancestral groups (K) ranging from 1 to 10 was tested, using a Monte-Carlo Markov Chain (MCMC) with 15,000 iterations (10,000 burn-in). To assess the consistency of the results, 10 runs per K value were performed, each one considering an admixture model with uncorrelated allele frequencies among populations. The  $\Delta K$  criteria implemented in STRUCTURE HARVESTER<sup>66</sup> v. 0.6.94 (July 2014) was used to estimate the most likely number of genetic groups. Since the most likely K value was 2, STRUCTURE was re-run with a fixed ALPHA value of 0.5 (*i.e.* 1/K) in order to minimize the effect of bias in numerical consistency of the two population subsets, as suggested by Wang<sup>67</sup>.

A Principal Component Analysis (PCA) was focused on estimating admixture between cultivated and wild *vinifera* accessions. To mitigate the bias due to unbalanced ancestral populations, we used a two-step approach consisting in (1) generating principal components (PCs) using balanced ancestral populations, (2) projecting the rest of samples on those pre-calculated PCs (see e.g.<sup>68</sup>). Equal size samples (N = 60) of ancestral cultivated and wild accessions were selected randomly, after excluding Lambrusco-named varieties and putatively admixed accessions based on prior information (varieties previously reported as genetically close to *sylvestris* and wild accessions with flower morphology intermediate between female and hermaphrodite).

Principal components were computed using the *prcomp* function from the *stats* R package (R Core Team, 2022). All the samples not included in ancestral populations were then projected on those PCs. Hierarchical clustering was performed by using the simple matching dissimilarity and the Unweighted Neighbour Joining (UwNJ) method implemented in DARwin v. 6<sup>69</sup>.

### Parentage analysis

Parent-offspring relationships were investigated by a likelihood-based method using the software Cervus v. 3.0.7<sup>70</sup>, accounting for a genotyping error rate of 0.01 in likelihood calculations. For each parent-offspring duos a LOD score (natural log of the overall likelihood ratio) is reported.

### Data availability

All data generated or analysed during this study are included in this published article and its supplementary information files.

Received: 16 November 2023; Accepted: 21 May 2024

Published online: 29 May 2024

### References

- Zohary, D. *The domestication of the grapevine Vitis vinifera L. in the Near East*, in (ed. McGovern, P., Fleming, S.J. & Katz, S.) *The origin and ancient history of wine* (Routledge, London). Reprinted *Food & Nutrition in History & Anthropology Series 11*. ISBN 90-5699-552-9 (2004).
- This, P., Lacombe, T. & Thomas, M. R. Historical origin and genetic diversity of wine grapes. *Trends Genet.* **22**(9), 511–519. <https://doi.org/10.1016/j.tig.2006.07.008> (2006).
- McGovern, P., Fleming, S. J. & Katz, S. *The origin and ancient history of wine* (Routledge, London). Reprinted *Food & Nutrition in History & Anthropology Series, 11*. ISBN 90-5699-552-9 (2004).
- Zhou, Y., Massonnet, M., Sanjak, J. S., Cantù, D. & Gaut, B. S. Evolutionary genomics of grape (*Vitis vinifera* ssp. *vinifera*) domestication. *Proc. Natl. Acad. Sci. U.S.A.* **114**(44), 11715–11720. <https://doi.org/10.1073/pnas.1709257114> (2017).
- Zhou, Y., Muyle, A. & Gaut, B. S. *Evolutionary Genomics and the Domestication of Grapes*, in (eds. Cantù, D., Walker, M.) *The Grape Genome* (Compendium of Plant Genomes, Springer, Cham.) [https://doi.org/10.1007/978-3-030-18601-2\\_3](https://doi.org/10.1007/978-3-030-18601-2_3) (2019).
- Dong, Y. et al. Dual domestications and origin of traits in grapevine evolution. *Science* **379**, 892–901. <https://doi.org/10.1126/science.add8655> (2023).
- Olmo, H. P. *The origin and domestication of the vinifera grape*, in (eds. McGovern, P., Fleming, S. J. & Katz, S.) *The Origin and Ancient History of Wine* (Routledge, London). Reprinted *Food & Nutrition in History & Anthropology Series 11*. ISBN 90-5699-552-9 (2004).
- Myles, S. et al. Genetic structure and domestication history of the grape. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 3530–3535. <https://doi.org/10.1073/pnas.1009363108> (2011).

9. De Andrés, M. T. *et al.* Genetic diversity of wild grapevine populations in Spain and their genetic relationships with cultivated grapevines. *Mol. Ecol.* **21**(4), 800–816. <https://doi.org/10.1111/j.1365-294X.2011.05395.x> (2012).
10. Riaz, S. *et al.* Genetic diversity analysis of cultivated and wild grapevine (*Vitis vinifera* L.) accessions around the Mediterranean basin and Central Asia. *BMC Plant Biol.* **18**, 137. <https://doi.org/10.1186/s12870-018-1351-0> (2018).
11. De Michele, R. *et al.* Phylogenetic relationship among wild and cultivated grapevine in Sicily: A hotspot in the middle of the Mediterranean Basin. *Front. Plant Sci.* <https://doi.org/10.3389/fpls.2019.01506> (2019).
12. Magris, G. *et al.* The genomes of 204 *Vitis vinifera* accessions reveal the origin of European wine grapes. *Nat. Commun.* **12**, 7240. <https://doi.org/10.1038/s41467-021-27487-y> (2021).
13. Arroyo-García, R. *et al.* Multiple origins of cultivated grapevine (*Vitis vinifera* L. ssp. *sativa*) based on chloroplast DNA polymorphisms. *Mol. Ecol.* **15**(12), 3707–3714. <https://doi.org/10.1111/j.1365-294X.2006.03049.x> (2006).
14. Cunha, J. *et al.* Genetic relationships among Portuguese cultivated and wild *Vitis vinifera* L. Germplasm. *Front. Plant Sci.* **11**, 127. <https://doi.org/10.3389/fpls.2020.00127> (2020).
15. Grassi, F. & De Lorenzis, G. Back to the origins: Background and perspectives of grapevine domestication. *Int. J. Mol. Sci.* **22**, 4518. <https://doi.org/10.3390/ijms22094518> (2021).
16. D’Onofrio, C. Introgression among cultivated and wild grapevine in Tuscany. *Front. Plant Sci.* <https://doi.org/10.3389/fpls.2020.00202> (2020).
17. Zecca, G. *et al.* Molecular grapevine: Sylvestris, hybrids or cultivars that escaped from vineyards? Molecular evidences in Sardinia. *Plant Biol.* **12**(3), 155–162. <https://doi.org/10.1111/j.1438-8677.2009.00226.x> (2010).
18. Bowers, J. *et al.* Historical genetics: The parentage of Chardonnay, Gamay, and other wine grapes of Northeastern France. *Science* **285**(5433), 1562–1565. <https://doi.org/10.1126/science.285.5433.1562> (1999).
19. Lacombe, T. *et al.* Large-scale parentage analysis in an extended set of grapevine cultivars (*Vitis vinifera* L.). *Theor. Appl. Genet.* **126**, 401–414. <https://doi.org/10.1007/s00122-012-1988-2> (2013).
20. Bacilieri, R. *et al.* Genetic structure in cultivated grapevines is linked to geography and human selection. *BMC Plant Biol.* **13**, 25. <https://doi.org/10.1186/1471-2229-13-25> (2013).
21. Cunha, J. *et al.* Grapevine cultivar “Alfrocheiro” or “Bruñal” plays a primary role in the relationship among Iberian grapevines. *Vitis* **54**, 59–65 (2015).
22. Laucou, V. *et al.* Extended diversity analysis of cultivated grapevine *Vitis vinifera* with 10K genome-wide SNPs. *PLoS One* **13**(2), e0192540. <https://doi.org/10.1371/journal.pone.0192540> (2018).
23. Raimondi, S. *et al.* DNA-based genealogy reconstruction of Nebbiolo, Barbera and other ancient grapevine cultivars from north-western Italy. *Sci. Rep.* **10**, 15782. <https://doi.org/10.1038/s41598-020-72799-6> (2020).
24. Maraš, V. *et al.* Population genetic analysis in old Montenegrin vineyards reveals ancient ways currently active to generate diversity in *Vitis vinifera*. *Sci. Rep.* **10**, 15000. <https://doi.org/10.1038/s41598-020-71918-7> (2020).
25. D’Onofrio, C. *et al.* Parentage atlas of Italian grapevine varieties as inferred from SNP genotyping. *Front. Plant Sci.* <https://doi.org/10.3389/fpls.2020.605934> (2021).
26. Miller, A. J. & Gross, B. L. From forest to field: Perennial fruit crop domestication. *Am. J. Bot.* **98**(9), 1389–1414. <https://doi.org/10.3732/ajb.1000522> (2011).
27. Ramos-Madriral, J. *et al.* Palaeogenomic insights into the origins of French grapevine diversity. *Nat. Plants* **5**, 595–603. <https://doi.org/10.1038/s41477-019-0437-5> (2019).
28. Arnold, C., Gillet, F. & Gobat, J. M. Occurrence of the wild vine *Vitis vinifera* ssp. *sylvestris* in Europe. *Vitis* **37**, 159–170. <https://doi.org/10.5073/vitis.1998.37.159-170> (1998).
29. Ocete, R. *et al.* Characterization of *Vitis vinifera* L. subspecies *sylvestris* (Gmelin) Hegi in the Ebro river Basin (Spain). *Vitis* **50**, 11–16 (2011).
30. Arroyo-García, R. A. & Revilla, E. The Current Status of Wild Grapevine Populations (*Vitis vinifera* ssp. *sylvestris*) in the Mediterranean Basin, In *The Mediterranean genetic code—Grapevine and olive* (Sladonja Barbara, Intech). <https://doi.org/10.5772/52933> (2013).
31. Doulati-Baneh, H. *et al.* Genetic characterization of some wild grape populations (*Vitis vinifera* subsp. *sylvestris*) of Zagros mountains (Iran) to identify a conservation strategy. *Plant Genet. Resour.* **13**(1), 27–35. <https://doi.org/10.1017/S1479262114000598> (2015).
32. Zdunić, G. *et al.* Genetic structure and relationships among wild and cultivated grapevines from Central Europe and part of the Western Balkan Peninsula. *Genes* **11**(9), 962. <https://doi.org/10.3390/genes11090962> (2020).
33. Failla, O., Anzani, R. & Scienza, A. L. vite selvatica in Italia: Diffusione, caratteristiche e conservazione del germoplasma. *Vignevini* **1**(2), 37–46 (1992).
34. Biagini, B. *et al.* Italian wild grapevine (*Vitis vinifera* L. subsp. *sylvestris*) population: Insights into eco-geographical aspects and genetic structure. *Tree Genet. Genomes* **10**, 1369–1385. <https://doi.org/10.1007/s11295-014-0767-4> (2014).
35. Garfi, G. *et al.* Habitat features and genetic integrity of wild grapevine *Vitis vinifera* L. subsp. *sylvestris* (C.C. Gmel.) Hegi populations: A case study from Sicily. *Flora* **208**, 538–548. <https://doi.org/10.1016/j.flora.2013.08.005> (2013).
36. Sereni, E. *Terra nuova e buoi rossi* (Einaudi, 1981).
37. De Crescentiis, P. 1309. *Opus ruralium commodorum* (ed. Strasbourg 1486; reprint by Biblioteca internazionale “La vigna”, Vicenza, Italy, 2010).
38. Levadoux, L. Les populations sauvages et cultivées de *Vitis vinifera* L. *Annales de l’Amélioration des Plantes* **1**, 59–118 (1956).
39. Grassi, F. *et al.* Evidence of a secondary grapevine domestication centre detected by SSR analysis. *Theor. Appl. Genet.* **107**, 1315–1320 (2003).
40. Lacombe, T. *Contribution à l’étude de l’histoire évolutive de la vigne cultivée (Vitis vinifera L.) par l’analyse de la diversité génétique neutre et de gènes d’intérêt* (Doctoral thesis, Montpellier SupAgro—Centre International d’Etudes Supérieures en Sciences Agronomiques, Montpellier, France, 2012).
41. Boccacci, P., Torello Marinoni, D., Gambino, G., Botta, R. & Schneider, A. Genetic characterization of endangered grape cultivars of Reggio Emilia province. *Am. J. Enol. Vitic.* **56**(4), 411–416 (2005).
42. Bowers, J. E. & Meredith, C. P. The parentage of a classic wine grape, Cabernet Sauvignon. *Nat. Genet.* **16**, 84–87 (1997).
43. Di Vecchi Staraz, M. *et al.* Genetic structuring and parentage analysis for evolutionary studies in grapevine: Kin group and origin of the cultivar Sangiovese revealed. *J. Am. Soc. Hortic. Sci.* **132**(4), 514–524 (2007).
44. Bowers, J. E., Siret, R., This, P., Boursiquot, J. M. & Meredith, C. P. A single pair of parents proposed for a group of grapevine varieties in Northeastern France. *Acta Hort.* **528**, 129–132. <https://doi.org/10.17660/ActaHortic.2000.528.15> (2000).
45. Regner, F., Stadlbauer, A. & Eisenheld, C. Heunisch x Fränkisch, ein wichtiger Genpool europäischer Rebsorten (*Vitis vinifera* L. *sativa*). *Wein-Wissenschaft* **53**(3), 114–118 (1998).
46. Crespan, M. & Milani, N. The Muscats: A molecular analysis of synonyms, homonyms and genetic relationships within a large family of grapevine cultivars. *Vitis* **40**, 23–30 (2001).
47. Péros, J.-P., Berger, G., Portemont, A., Boursiquot, J.-M. & Lacombe, T. Genetic variation and biogeography of the disjunct *Vitis* subg. *Vitis* (Vitaceae). *J. Biogeogr.* **38**, 471–486. <https://doi.org/10.1111/j.1365-2699.2010.02410.x> (2011).
48. Margaryan, K. *et al.* Grapevine genetic resources of Armenia: Molecular fingerprinting and phylogenetic relationships among wild and cultivated grapevine. *Vitis* **62**, 11–22. <https://doi.org/10.5073/vitis.2023.62.special-issue.11-22> (2023).

49. Vecchi-Staraz, Di. *et al.* Low level of pollen-mediated gene flow from cultivated to wild grapevine: Consequences for the evolution of the endangered subspecies *Vitis vinifera* L. subsp. *silvestris*. *J. Hered.* **100**(1), 66–75. <https://doi.org/10.1093/jhered/esn084> (2009).
50. Meléndez, E. *et al.* Evolution of wild and feral vines from the Ega river gallery forest (Basque country and Navarra, Spain) from 1995 to 2015. *OENO One* <https://doi.org/10.20870/oeno-one.2016.50.2.782> (2016).
51. Marrano, A., Micheletti, D., Lorenzi, S., Neale, D. & Grando, M. S. Genomic signatures of different adaptations to environmental stimuli between wild and cultivated *Vitis vinifera* L. *Hortic. Res.* <https://doi.org/10.1038/s41438-018-0041-2> (2018).
52. Freitas, S. *et al.* Pervasive hybridization with local wild relatives in Western European grapevine varieties. *Sci. Adv.* <https://doi.org/10.1126/sciadv.abi8584> (2021).
53. Bronner, J. P. *Die wilden Trauben der Rheinthalen* (Mohr, 1857).
54. Robinson, J., Hardy, J. & Vouillamoz, J. *Winegrapes* (Penguin books, 2012).
55. Tanara, V. *L'economia del cittadino in villa* (1644) (G.B. and G. Coruo, Rome, Reprint 1651).
56. Giusti, G. *Rosso rubino, profumo di viola: appunti per una ricerca sul lambrusco* (Futura libri, 1976).
57. Cazzola, F. Foreste artificiali. Espansione e declino della piantata padana (sec. XV–XX), in *I frutti di Demetra*, 23–30 (2007).
58. De Lattin, G. Über der Ursprung und die Verbreitung der Reben. *Der Zuchter* **11**(8), 217–225 (1939).
59. Gallo, A. *Le dieci giornate della vera agricoltura e piaceri della villa* (G.B. Bozzola, Brescia, 1564).
60. Meglioni, S. *et al.* Conoscere per tutelare il patrimonio viticolo: l'esperienza della provincia di Reggio Emilia. *L'Informatore Agrario* **23**, 2–6 (2013).
61. Schneider, A. *et al.* Identification and characterization of *Vitis vinifera* subsp. *silvestris* populations in north western Italy. *Vitis* **54**, 223–225 (2015).
62. Lazzari, G., Merloni, N. & Saiani, D. *Flora Pinete storiche di Ravenna—San Vitale, Classe, Cervia. Parco Delta del Po* (Regione Emilia-Romagna. Quaderni IBIS, Inventari della Biodiversità Specifica, Flora di Ravenna). <https://doi.org/10.13140/2.1.3083.6007> (2010).
63. Thomas, M. R., Matsumoto, S., Cain, P. & Scott, N. S. Repetitive DNA of grapevine: Classes present and sequences suitable for cultivar identification. *Theor. Appl. Genet.* **86**, 173–180 (1993).
64. Ruffa, P., Raimondi, S., Boccacci, P., Abbà, S. & Schneider, A. The key role of “Moscato bianco” and “Malvasia aromatica di Parma” in the parentage of traditional aromatic grape varieties. *Tree Genet. Genomes* **12**(3), 1–14. <https://doi.org/10.1007/s11295-016-1006-y> (2016).
65. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–995 (2000).
66. Earl, D. A. & vonHoldt, B. M. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**(2), 359–361. <https://doi.org/10.1007/s12686-011-9548-7> (2012).
67. Wang, J. The computer program STRUCTURE for assigning individuals to populations: Easy to use but easier to misuse. *Mol. Ecol. Resour.* **17**, 981–990. <https://doi.org/10.1111/1755-0998.12650> (2017).
68. Migicovsky, Z. *et al.* Genomic ancestry estimation quantifies use of wild species in grape breeding. *BMC Genom.* **17**, 478. <https://doi.org/10.1186/s12864-016-2834-8> (2016).
69. Perrier, X. & Jacquemoud-Collet, J. P. DARwin software <https://darwin.cirad.fr/> (2006).
70. Kalinowski, S. T., Taper, M. L. & Marshall, T. C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* **16**, 1099–1106. <https://doi.org/10.1111/j.1365-294x.2007.03089.x> (2007).

## Acknowledgements

We are grateful to Thierry Lacombe and Jean-Michel Boursiquot for the fruitful discussion during the survey on the *silvestris* populations in France. We also thank Mauro Catena for the shared information on Lambruscos from Emilia-Romagna. Amedeo Moine is also acknowledged for his support in several chloroplastic marker analyses.

## Author contributions

AS and SR conceived the study, collected and preserved the investigated genetic resources. PR and PB performed DNA extraction and genotyping. MF provided historical and ecological information on some wild populations, Lambruscos and other key grape cultivars. SR and GT took care of data curation and genetic statistical analyses. AS wrote the original draft that was critically revised by SR and GT. All authors read and approved the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-62774-w>.

**Correspondence** and requests for materials should be addressed to A.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024