EVOLUTION, DOMESTICATION AND DIVERSITY OF THE CULTIVATED SPECIES OF THE VITACEAE FAMILY IN THE CARPATHIAN BASIN FROM THE POINT OF MORPHOMETRIC ANALYSES

ZOLTÁN MRAVCSIK

Gödöllő

2019
<table>
<thead>
<tr>
<th>Name of the doctoral school:</th>
<th>Doctoral school of environmental sciences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branch of science:</td>
<td>Environmental sciences</td>
</tr>
<tr>
<td>Head:</td>
<td>Dr. Erika Csákiné Michéli</td>
</tr>
<tr>
<td></td>
<td>Professor, head of institute</td>
</tr>
<tr>
<td></td>
<td>Szent István University,</td>
</tr>
<tr>
<td></td>
<td>Faculty of Agricultural and Environmental Sciences,</td>
</tr>
<tr>
<td></td>
<td>Institute of Environmental Sciences,</td>
</tr>
<tr>
<td></td>
<td>Department of Soil Science and Agricultural Chemistry</td>
</tr>
<tr>
<td>Supervisors:</td>
<td>Dr. Ferenc Gyulai</td>
</tr>
<tr>
<td></td>
<td>Professor, doctor of the Hungarian Academy of Sciences</td>
</tr>
<tr>
<td></td>
<td>Szent István University,</td>
</tr>
<tr>
<td></td>
<td>Faculty of Agricultural and Environmental Sciences,</td>
</tr>
<tr>
<td></td>
<td>Institute of Nature Conservation and</td>
</tr>
<tr>
<td></td>
<td>Environmental Management</td>
</tr>
<tr>
<td></td>
<td>Department of Nature Conservation and</td>
</tr>
<tr>
<td></td>
<td>Landscape Ecology</td>
</tr>
<tr>
<td></td>
<td>Dr. József Berke</td>
</tr>
<tr>
<td></td>
<td>Professor, candidate of the Hungarian Academy of Sciences</td>
</tr>
<tr>
<td></td>
<td>Dennis Gabor College, Institute of Informatics</td>
</tr>
</tbody>
</table>

Approval of the head of the doctoral school

Approval of the supervisors

........................................................

........................................................

........................................................

........................................................
1. Introduction and objectives

Although vine and wine culture have rich traditions in Hungary, we only have limited information on the starting of viticulture and the development of the different varieties. There are many questions to be answered regarding the domestication process of vine and the history of viticulture. We do not know exactly which factors lead to the current great genetic diversity of the species of the Vitis genus, especially in the case of wine grape (Vitis vinifera L.). The origin of the old grape varieties is also obscure, and usually the historical appearance of the different varieties can be only approximated based on unreliable sources.

Archaeological findings carry important information regarding the past. After accurate morphological identification, archaeobotanical research can be used for the reconstruction of the natural and cultivated vegetation of the past ages. Regarding the applied methods and tools, technical development enables for the elaboration of new methodologies leading to more accurate and faster identification. Using the advantages of digital technologies is also beneficial in the case of vine, since the plant remains of the old ages keep turning up, and their analysis provides new data by which our current knowledge can become more accurate.

It is assumed that wild grape (Vitis sylvestris C. C. Gmel.) was the ancestor of wine grape, and several evidences have been found confirming its use in the Carpathian Basin among which seed remains form an exceptional category. The seeds of the two species can be differentiated based on their morphological traits, and they also carry the signs of domestication.

Seed analyses can be helpful also in other tasks. During my further investigations I experienced that the vegetation of abandoned vineyards still contains some old grape varieties. Although at some places cultivated plant species are dominant, many old grape and fruit varieties can be found in the vineyards even today. Since the recognition of the genetic value of the old Hungarian varieties and landraces there is a growing demand aiming at their saving. These initiatives are hindered by the fact that in many cases valuable varieties are difficult to differentiate from direct producing grape varieties becoming so widespread in the 20th century.

Wild grape (Vitis sylvestris C. C. Gmel.) is a protected species in Hungary, because its natural stands became scarce, and their survival is threatened today. Besides the disappearing and the transformation of its habitats, invasive Vitis taxa (Vitis riparia Michx., Vitis labrusca L. and Vitis rupestris Schee.) also considerably endanger its stands, because by running wild they easily occupy habitats, and due to their endurance they crowd out the indigenous species. In addition, pure wild grape stands may become genetically contaminated, or may disappear due to the process of hybridization.
During my research I set the following objectives:

- To collect information on the origin of the different grape varieties.
- To develop a seed collection of *Vitis* species, which can be the basis for morphometric analyses.
- To elaborate an appropriate method for the registration of morphometric traits by which the typical characteristics of the seeds can be rapidly measured and stored as data, enabling for further processing. I tried to implement this by using easily accessible devices making the process cost-effective.
- To find those seed related parameters – after the development of the database – by which wine grape (*Vitis vinifera* L.) and its wild ancestor (*Vitis sylvestris* C.C. Gmel.) can be distinguished. This can be the basis for computerized sorting leading to faster and more accurate data processing. As a result, digital morphometric methods could be used even for the identification of varieties or groups of varieties.
- To investigate if the different varieties of *Vitis vinifera* can be distinguished from the direct producing varieties and whether wild grape can be differentiated from the endangering *Vitis* taxons by using the methods of seed morphometry.
- To investigate if similarly to wine grape the signs of domestication – which can be characterized through morphologic and dimensional traits – can be witnessed also in the case of other vine species or not.
2. Material and method

2.1. Development of a grape seed collection

One of the main objectives of my investigations was the development of a database, which could be used during the identification of grape seed remains. Since seeds were the subject of my analyses, it was inevitable to obtain the seeds of certain *Vitis* taxons. As there are no grape seed collections including recent grape varieties or other vine species in Hungary, first I had to develop one.

Based on the available sources and the instructions of the Hungarian specialists of this field, I collected the seeds of old varieties from variety collections. Besides, other species and varieties also have been collected. Due to the analyses focusing on the utilization of *Vitis sylvestris* C. C. Gmel. and *Vitis vinifera* L. and the domestication of vine, wild grape samples also needed to be archived. Since there is a growing demand for old fruit varieties and landraces, whose specimens running wild have been found in abandoned vineyards (Mravcsik et al. 2009), seed based identification can be helpful also during their recognition. For this end, it was necessary to store also the seeds of direct producing varieties to see if they can be differentiated from the varieties of *Vitis vinifera* L. based on their morphological traits. The analysis of the morphometric traits of the seeds of invasive *Vitis* species was an important issue because of the protection of the decreasing wild grape populations; therefore other species also have been included in the collection. Since it was important to process authentic samples, I primarily processed those ones originating from variety collections.

These were the following:

- University of Pécs (PTE), Research Institute of Viticulture and Oenology (Pécs),
- National Agricultural Research and Innovation Centre (NAIK), Research Institute of Viticulture and Oenology, Badacsony Research Station (Badacsonytomaj),
- Tokaj Research Institute, Viticulture and Oenology Research Nonprofit Ltd. (and its legal predecessor) (Tarcal),
- Eszterházy Károly University (former KRF), Research Institute of Viticulture and Oenology (Eger),
- Pannon University, Georgikon Faculty, collections of the Department of Horticulture (Keszthely, Cserszegtomaj),
- Collection of the Patricius winery (Bodrogkisfalud, Várhegy vineyard).
During the collection of grape species and varieties I tried to gather as many samples as possible, in order to develop a large collection and database. Only seeds from ripe berries have been selected, and not only the largest seeds but also the fully developed smaller ones have been collected. The reason for that was to have a collection fully representing the characteristic seed populations of the specific varieties. During collection I selected approximately 150-250 seeds per variety, for which usually 3-6 bunches have been used. In the case of some invasive *Vitis* species only 1-2 bunches could be harvested due to the low yield levels. The different samples were put in separate plastic bags with a piece of paper recording the name of the species or the variety. After the collection and the delivery of the samples, selected seeds have been cleaned, dried, and finally they got stored in airtight containers.

This collection is gap filling and unique, and it provides several opportunities for further research and a solid basis for future seed morphology analyses. Also it can be used as an easily accessible source enabling for long term storage in molecular genetics analyses – representing the most reliable tool for variety identification today –, although these processes require the destruction of the seed. However, in the case of developing a DNA library the necessary data can be stored after a single processing phase.

### 2.2. Introduction of the used methods and tools

Since I used Fovea Pro 4.0 software for the morphometric analyses, it had a significant impact on the methodology itself (Russ 2006). The first step included taking images of the different *Vitis* sp. seeds, during which I tried to make a good quality basis for further data processing, because it may considerably affect the whole process. Images can be taken by using various devices, and I also tested more of them during my work. As a general rule, it can be stated that mistakes should be avoided during taking images, because it is time-consuming and difficult to eliminate them later during post-processing. In my case it was a prerequisite to keep the same distance between the camera and the different samples, and to fix the optical axle of the camera at right angles to the plane of the samples during taking the images. Besides, applying the same illumination conditions was also required. Factors increasing the efficiency and the successfulness of the method included recording the largest possible number of objects at the same time, taking images with the highest possible resolution, minimizing the occurrence of noises and unwanted objects (shading and dust) and avoiding the contact of the different elements (seeds). I tried to meet these expectations as much as possible.

I started my analyses by using a microscope, because it is essential during the archaeobotanical identification of the seeds, and I thought that the most detailed and the highest magnification images can be made with this device. However, I faced several negative features of
this tool during my research. I could put only two-three grape seeds on the object slide in a way that they still fall within the view of the camera mounted on the microscope, but in spite of the round illumination significant shading has been witnessed on the images. Furthermore, it was problematic to find an appropriate focus distance due to the differences of the seeds, and setting the depth of focus was also difficult to set even in the case of a single seed.

Taking images through a microscope and performing calibrated measuring with the related software are suitable only for the analysis of a few parameters of a limited number of samples in case excellent image definition for every detail of the object is not a prerequisite, or if the object is only two-dimensional. As I wanted to elaborate a method which is accessible for everyone, I had to look for another device also because of the high price of the microscope.

Scanners enable for taking images of a large number of samples simultaneously within a short period of time, in addition they are affordable. When taking the images I covered the seeds with a white sheet of paper in order to reach the highest quality of seed silhouettes, but the result was not perfect this way either since shades appeared among the seeds, which needed to be removed during post-processing. In order to reliably detect differences among objects that seem almost the same, one shall try to take the best “possible” and the most detailed digital images, which can be implemented by using a camera.

Similarly to scanners, cameras can be used for taking high resolution images of a large number of samples. During my analyses I used a **Nikon D5200** digital camera equipped with the following accessories:

- AF-S DX Micro NIKKOR 40mm f/2.8G objective
- Hoya HMC UV(C) filter
- Meike ML-L3 infrared remote trigger (Nikon)
- SanDisk SDHC 8GB Ultra (class 10) memory card
- Manfrotto 190XPROB tripod, 056 3D tripod head
- Dörr Daf-14 circular flash (Nikon)
- Air-level

When taking the images I placed a sheet of paper under the seeds providing a white background this way. Grape seeds were put on this one-by-one with their backside facing down, and lying on their hilums (*chalaza*). In this position the total width and length of the seeds and their morphological traits could be accurately recorded. All digital images have been taken by applying the same camera settings.
2.3. Image processing and measuring

Raw images are not suitable for performing measurements, since they contain several unwanted elements (dust and other “noise”), which need to be mitigated. I cleaned the images by using the Adobe Photoshop CS2 program.

Measurements were taken by using the saved TIFF files with the help of the Fovea Pro 4.0 program embedded into the Adobe Photoshop CS2. In total, 56 parameters of 6734 seeds have been recorded, which were exported to the Microsoft Excel 2016 spreadsheet program. During data assessment I took 34 indices into account, including dimensional (e.g. area, length, width, radius of the circles which can be drawn inside and around the seed), morphological (e.g. form factor, roundness, aspect ratio, convexity, symmetry) and colour and intensity related (mean red, green and blue colour values, tone, chroma) parameters.

2.4. Assessment of the results

During the assessment of the results I used the CHAID (Chi-squared Automatic Interaction Detector) method (decision making tree method) by applying the IBM SPSS v.25 program package for the differentiation of the seeds of wild grape and wine grape. Furthermore, I run logistic regression and principal component analysis in case of all the investigated seeds for analysing how to distinguish the different groups (wine grape, wild grape, other Vitis species and direct producing varieties). Besides, I also made assessments by using the Microsoft Excel 2016 program for the visualization of larger data sets. Through data display much information can be obtained helping us to reveal similarities and relationships.
3. Results

3.1. Development of a seed collection

By using seeds obtained from the variety collections I developed a collection consisting of 172 items with the following composition:

- 137 pieces of *Vitis vinifera* L. items, including 115 different varieties,
- 22 seed populations of 16 pieces of other *Vitis* species, including 5 *Vitis sylvestris* items,
- 13 seed populations of 7 pieces of direct producing grape varieties, including “Baco” and “Piros delaware” being the progeny of *Vitis vinifera* L. as opposed to the remaining varieties originating from other vine species.

3.2. Development of a morphometric database of seeds

So far the seeds of 42 items have been processed including 32 old *Vitis vinifera* L. varieties, 4 direct producing varieties and 6 species belonging to the *Vitis* genus (from which 2 items are *Vitis sylvestris* C.C.Gmel) (see Table 1). The database also contains the parameters of the images taken with the help of the above mentioned set of tools, which were measured by using the FOVEA Pro 4.0 software.

**Table 1: Measured grape species and varieties included in the morphometric database of seeds**

<table>
<thead>
<tr>
<th>Number</th>
<th>Number and name of the item</th>
<th>Number</th>
<th>Number and name of the item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>40.Csókaszőlő</td>
<td>22.</td>
<td>126.Merlot</td>
</tr>
<tr>
<td>2.</td>
<td>46.Vitis amurensis</td>
<td>23.</td>
<td>127.Kéknyelű</td>
</tr>
<tr>
<td>3.</td>
<td>47.Vitis labrusca L.</td>
<td>24.</td>
<td>130.Szürkebarát</td>
</tr>
<tr>
<td>4.</td>
<td>52.Vitis riparia</td>
<td>25.</td>
<td>134.Betyárszőlő</td>
</tr>
<tr>
<td>5.</td>
<td>54.Vitis rupestris</td>
<td>26.</td>
<td>143.Rozaki</td>
</tr>
<tr>
<td>6.</td>
<td>60.Heunisch</td>
<td>27.</td>
<td>147.Ál-kék hajnos</td>
</tr>
<tr>
<td>8.</td>
<td>82.Tűskéspúpú</td>
<td>29.</td>
<td>159.Sárfehér</td>
</tr>
<tr>
<td>9.</td>
<td>84.Lisztesfehér</td>
<td>30.</td>
<td>160.Kecskecsőcsű</td>
</tr>
<tr>
<td>10.</td>
<td>94.Ezerjó</td>
<td>31.</td>
<td>166.Gohér</td>
</tr>
<tr>
<td>12.</td>
<td>104.Fehér gohér</td>
<td>33.</td>
<td>168.Sárgamuskotály</td>
</tr>
<tr>
<td>13.</td>
<td>108.Fehér szlanka</td>
<td>34.</td>
<td>169.Furmint</td>
</tr>
<tr>
<td>14.</td>
<td>112.Gyöngyfehér</td>
<td>35.</td>
<td>172.Purcsín</td>
</tr>
<tr>
<td>15.</td>
<td>113.Bajor kék</td>
<td>36.</td>
<td>187.Fehér járdovány</td>
</tr>
<tr>
<td>16.</td>
<td>117.Kadarka</td>
<td>37.</td>
<td>196.Elvira</td>
</tr>
<tr>
<td>17.</td>
<td>119.Rajnai rizling</td>
<td>38.</td>
<td>197.Izabella</td>
</tr>
<tr>
<td>18.</td>
<td>121.Cab.sauvignon</td>
<td>39.</td>
<td>198.Concord</td>
</tr>
<tr>
<td>19.</td>
<td>122.Cab.franc</td>
<td>40.</td>
<td>199.Piros delaware</td>
</tr>
<tr>
<td>20.</td>
<td>123.Juhfark</td>
<td>41.</td>
<td>216. Vitis sylvestris S-6/1</td>
</tr>
<tr>
<td>21.</td>
<td>125.Bakator</td>
<td>42.</td>
<td>218. Vitis sylvestris S-6/2</td>
</tr>
</tbody>
</table>
Altogether, 56 parameters of 6734 grape seeds have been included in the database. In addition, the minimum, maximum and mean values of the different items also have been calculated and stored in a separate document.

3.3. Digital image processing methods for the species and varieties of the *Vitis* genus

I elaborated a methodology accessible for everyone by which the seeds of any grape species or varieties can be recorded and measured – even in the case of a larger number of samples –, meaning that they can be compared to the specific parameters already available in the database.

**When applying the method the following steps needed to be taken:**

1. Putting the grape seeds on the white sheet of paper functioning as the background and the plotting paper fixed under that.
2. Joining the elements of the digital camera (objective, circular flash).
3. Fixing the device on the tripod, and setting it accurately above the object to be analysed.
4. Fixing the optical axle of the camera at right angles to the plane of the samples with the help of the air-level and the spirit level.
5. Performing settings by applying: f/8 shutter, 1/125 seconds of exposure time, 100 ISO value and 40 mm of focus distance.
6. Taking images in a dark room with the help of the circular flash and the remote trigger.
7. Saving images in RAW and TIFF formats with 48 bit depth.

**Image processing**

1. Opening the images file with the Photoshop program.
2. Selecting seeds with the “Magicwand” tool.
3. Inverting selection, and deleting the background.
4. Saving the images in 48 bit TIFF format.
5. Opening the 48 bit image files with the program Fovea Pro 4.0.
6. Calibrating the images according to the plotting paper.
7. Performing measurements.

3.4. Assessment of morphometric measurements

During the processing of images taken with the help of a camera dimensional, morphological and colour related variables can be differentiated among the analysed parameters.
3.4.1. Dimensional parameters

Based on the dimensional indices of the analysed grape seeds it can be stated that the seeds of wild grape had the lowest values in the case of all the parameters (see Figure 1). The highest values have been recorded in the case of wine grape varieties, but for some traits (“width”, “equivalent diameter” and “inscribed radius”) the direct producing variety called “Concord” proved to be outstanding. If we assess the dimensional parameters of the two *Vitis sylvestris* C. C. Gmel. items together, it can be stated that they can be visually differentiated from the other species and varieties. When analysing a small number of samples, identification can be made more reliable in case two (or more) parameters are assessed in the same sample space.

In the case of seed length ampelographers set 4.5-5 mm as the limit value for the differentiation of *V. vinifera* and *V. sylvestris*. If the length of the seed is lower than this value, it is supposed to be a wild grape, while seeds exceeding this limit value belong to wine grape. (Schermann 1966). My results also verify this statement, which could be a good method for differentiating the two species. However, further factors also need to be analysed besides this one during the reliable differentiation of invasive *Vitis* species and wild grape.

Figure 1: Mean area values of the analysed grape seed items (wine grape: blue, other *Vitis* species: yellow, direct producing varieties: red, wild grape: green)
3.4.2. Morphological parameters

Morphological parameters include data related to the different ratio of dimensional values and the shape of the objects.

Morphological parameters are usually shape related traits, but many times they are values calculated from dimensional indices. In some cases (“Y-feret”, “roundness” and “elongation”) wild grape has the highest and lowest values, and also it represents one of the lowest values (1.36) for “aspect ratio”. The last three categories describe the elongation and the roundness of the seeds (“roundness”: difference from the shape of a circle, “elongation”: ratio of the area and the perimeter, “aspect ratio”: ratio of the length and the width). In the case of “roundness” its value reaches 0.70, meaning that the shape of this item is the closest to a circle (value of 1). The seeds of *Vitis sylvestris* C. C. Gmel. also reach low values (1.36 and 2.28) when assessing “aspect ratio” and “elongation”, indicating that they are not elongated but rather round, as it is also verified by the different sources of literature (Mangafa and Kotsakis 1996, Facsar 2000, Rivera et al. 2007, Gyulai 2009). The highest mean aspect ratio value (1.98) was measured in the case of the variety called “Gohér”, while the variety called “Heunisch” had the highest mean elongation value (4.12). The other *Vitis* species usually reach low values (“aspect ratio”: 1.27-1.40, “elongation”: 2.44-2.58) in the case of these indices.

It is also visible that morphological parameters are not suitable for the differentiation of *Vitis sylvestris* C. C. Gmel. and *Vitis* species occupying its natural habitats, therefore other values also need to be included in the assessment. However, it happens many times that wine grape varieties can be visually differentiated from the values of wild grape (e.g. in the case of Y-feret the difference is significant) and that of the other studied *Vitis* species, which is a clear sign of morphological differences.

3.4.3. Intensity and colour parameters

Colour and intensity related indices are useful when studying and comparing recent samples. In the case of archaeological samples they are not relevant, since these seeds already lost their original colour and texture. These indices are quite sensitive to the circumstances of taking images. Based on the results, it is clearly visible that the variety called “Sárfehér” has the highest measured values in the case of 6 indices (three colour channels, “brightness”, “minimum intensity” and “saturation”), while item No. 216 of wild grape has the lowest mean values in five cases. They do not differ significantly from the other items, but compared to each other and a few other items these parameters can be also useful during selection. It also must be noted that although the two varieties of *Vitis sylvestris* C. C. Gmel. are supposed to be genetically closely related, differences have been experienced between their colour parameters. In view of this and based on the data, it
seems that wild grape and wine grape cannot be differentiated by purely relying on them. In the case of “mean hue” it can be observed that the four direct producing varieties significantly differentiate from the other varieties (see Figure 2). Their average values range between 27.8-29.0, which is a significant difference compared to the other studied samples, therefore this is an aspect worth investigating during the separation of direct producing varieties.

Figure 2: Distribution of the mean hue values of the analysed items (wine grape: blue, other Vitis species: yellow, direct producing varieties: red, wild grape: green)

3.4.4. Differentiation of Vitis sylvestris C. C. Gmel. and Vitis vinifera L.

My analyses verified the findings of the former investigations, that is the size of the seeds increased, and the beak of the seed simultaneously elongated during the process of domestication (Schermann 1966, Jacquat and Martinoli 1996, Mangafa and Kotsakis 1996, Facsar 2000, Rivera et al. 2007, Gyulai 2009). There are several dimensional and morphological parameters verifying this during the morphometric comparison of the seeds of wild grape items and that of Vitis vinifera L. varieties, according to which the two species can be distinguished. Although dimensional data sets are suitable for the differentiation of the two species, morphological and colour and intensity related information are not so reliable. Therefore, it is advised to consider more parameters, especially when analyzing only a small number of samples, because according to the data, the seed populations of the different species or varieties may vary to a great extent.

In some cases Vitis sylvestris C. C. Gmel. and Vitis vinifera L. can be clearly distinguished based on the mean values of the measured parameters weighted by the standard deviation. Standard deviation is the squared average of the deviation of the specific values from the average, and shows
how the known values deviate from the average on the average (Závoti 2010). By taking into account this, my measurements found that the limit values of the following factors are characteristic of the two species, which can be also used as key parameters during identification (see Table 2).

The following parameters can be used for the differentiation of wild grape and wine grape:

Table 2: Limit values suitable for the differentiation of Vitis sylvestris C. C. Gmel. and Vitis vinifera L.

<table>
<thead>
<tr>
<th>Name of the parameter</th>
<th>Vitis sylvestris C. C. Gmel.</th>
<th>Vitis vinifera L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area (mm²)</td>
<td>≤ 9.30</td>
<td>≥ 13.19</td>
</tr>
<tr>
<td>Convex area (mm²)</td>
<td>≤ 9.64</td>
<td>≥ 13.81</td>
</tr>
<tr>
<td>Length (mm)</td>
<td>≤ 4.33</td>
<td>≥ 5.38</td>
</tr>
<tr>
<td>Width (mm)</td>
<td>≤ 3.11</td>
<td>≥ 3.34</td>
</tr>
<tr>
<td>Equivalent diameter (mm)</td>
<td>≤ 3.44</td>
<td>≥ 4.10</td>
</tr>
<tr>
<td>Inscribed radius (mm)</td>
<td>≤ 1.53</td>
<td>≥ 1.66</td>
</tr>
<tr>
<td>Y-feret</td>
<td>≤ 230.32</td>
<td>≥ 387.46</td>
</tr>
</tbody>
</table>

For the differentiation of wild grape and wine grape I developed a CHAID based decision making tree by using the SPSS v.25 software (Hámori 2001). This multivariate classification method separated the items into two groups by using three steps with an accuracy level of 99.52% (32 wrong and 6702 right classifications from 6734 cases). This also verifies that the two species can be reliably differentiated based on the measured parameters.

The discriminant analysis run by the SPSS v.25 program package also provided results suitable for differentiation (see Figure 3). In this analysis all the measured grape seeds have been visualized by two discriminant functions.
Based on the discriminant analysis it can be stated that wild grape items can be clearly differentiated from the other ones, and only a few seeds of the varieties called “Csókaszőlő” and ”Mézesféhér” belonging to the category of wine grape are closer to them. Direct producing varieties and those belonging to the category of wine grape do not separate from each other, but the analysed Vitis species, wild grape and the direct producing varieties form clearly separated groups, which are only slightly overlapping with the varieties of Vitis vinifera L. This result is also important from the point that wild grape and invasive grape species spreading on its natural habitats indicate some kind of separation through these functions. During the “logistic regression” analysis run by the IBM SPSS v.25 program I did not get any usable results due to the disturbing cross impacts experienced among the parameters. This drew my attention to the fact that some of the parameters analysed by FOVEA Pro 4.0 have similar characteristics, or there will be no difference between them in the case of grape seeds (e.g. ”area” – “convex area”). Also, there are more parameters which are calculated by using other measured parameters [e.g. “aspect ratio”: ratio of the length and the width]. This should be taken into account when performing further sensitive analyses, and some of the parameters need to be left out if necessary.

In the case of cultivated species genetic diversity usually decreases due to the impacts of domestication (selection), which can be also observed on the seeds and the fruits. Rovner and
Gyulai (2007) verified this phenomenon in the case of different plant species. They stated that the impacts of human selection can be observed in the diversity of the different parameters of the seeds irrespectively of the other environmental factors, and they get genetically fixed. During domestication a specific group of characteristics (size, shape or nutritional value) was the basis for seed selection, which decreased diversity, and lead to uniformity that can be witnessed in the unimodal shape of the distribution histograms calculated based on the parameters. It has been verified in the case of many domesticated plants and their progenitors that histograms developed from the measured data of the wild species had more peaks, while in the case of the domesticated species it had only one peak, and its shape was similar to a bell curve (Gyulai et al. 2015). The unimodal histogram can be exactly distinguished from the multimodal one by drawing trend lines. If the polynomial trend line no. 2 (dotted line on Figure 4-6) intersects the 2 periodic trend line of the moving mean (dashed line on Figure 4-6) more than twice, the histogram can be regarded multimodal, while in the case of only two intersection points the histogram is unimodal. I was curious what kind of trends could be observed in the case of wild grape and wine grape according to the seeds.

I analysed the two *Vitis sylvestris* C.C.Gmel. items separately and together as well, and I got similar results. In the case of the dimensional parameters I found that grape seeds are unevenly distributed between the maximum and the minimum values (see Figure 4 and 5).

![Figure 4: Multimodal distribution of the length values of wild grape (item No. 218)](image-url)
In order to reach the taxonomic level of *V. sylvestris*. C. Gmel., first I analyzed the two varieties of *Vitis vinifera* L. together. I did not get the expected unimodal histogram in this case either, since it was not similar to a Gauss curve, and it had more peaks. Regarding dimensional values, the two varieties together occupy a broader interval, but distribution is more uniform between the groups.

Since analysing all the studied wine grape varieties together would require accepting huge differences between the sample sizes compared to that of wild grape, I also analysed the varieties one-by-one visualizing the distribution of dimensional and morphological parameters. This time I also experienced the same, as I got multimodal histograms again (see Figure 6).
There can be several reasons for that, which require further research to reveal. It is likely that *Vitis vinifera* L. has a unique genetic diversity among the cultivated plants, which is confirmed by its more thousands of varieties. Wild grape is regarded to be the ancestor of wine grape, but we do not know exactly under which conditions did wine grape develop from the basic species, and if it can be regarded uniform or not. In addition, revealing differences among the geographical and ecological groups of varieties would be also important. Genetic studies dealing with this issue are recommended to be complemented with morphometric measurements in order to draw accurate conclusions.

### 3.4.5. Impacts of breeding on the morphological traits of the seeds of *Vitis labrusca* L.

In the case of wild grape the size of the seeds increased, and they also became elongated as a result of domestication (Schermann 1966, Jacquat and Martinoli 1996, Mangafa and Kotsakis 1996, Facsar 2000, Rivera *et al.* 2007, Gyulai 2009). The signs of convergence experienced in the field of cultivated plants can be also witnessed in the case of other species (see Table 3).

**Table 3: Comparison of some of the parameters of *Vitis labrusca* items and that of the variety called “Izabella”**

<table>
<thead>
<tr>
<th></th>
<th>Area (mm²)</th>
<th>Length (mm)</th>
<th>Equivalent diameter (mm)</th>
<th>Circumized radius (mm)</th>
<th>Perimeter (mm)</th>
<th>X-feret</th>
<th>Y-feret</th>
<th>Roundness</th>
<th>Aspect ratio</th>
<th>Elongation</th>
<th>Mean Hue</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Vitis labrusca</em></td>
<td>10.81</td>
<td>4.54</td>
<td>3.71</td>
<td>2.27</td>
<td>13.46</td>
<td>244.51</td>
<td>329.42</td>
<td>0.67</td>
<td>1.38</td>
<td>2.58</td>
<td>18.80</td>
</tr>
<tr>
<td>(No. 47)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Izabella</td>
<td>19.51</td>
<td>6.56</td>
<td>4.98</td>
<td>3.29</td>
<td>19.02</td>
<td>303.49</td>
<td>480.29</td>
<td>0.58</td>
<td>1.61</td>
<td>3.22</td>
<td>28.32</td>
</tr>
<tr>
<td>(No. 197)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td>8.70</td>
<td>2.02</td>
<td>1.27</td>
<td>1.02</td>
<td>5.56</td>
<td>58.98</td>
<td>150.87</td>
<td>0.09</td>
<td>0.23</td>
<td>0.64</td>
<td>9.52</td>
</tr>
<tr>
<td>Ratio of change (%)</td>
<td>80.45</td>
<td>44.56</td>
<td>34.32</td>
<td>44.90</td>
<td>41.28</td>
<td>24.12</td>
<td>45.80</td>
<td>13.92</td>
<td>16.68</td>
<td>24.84</td>
<td>50.65</td>
</tr>
</tbody>
</table>
The increase of seed size can be also verified by using the example of the invasive *Vitis labrusca* L. and its cultivated form, called “Izabella”. Their area almost doubled (from 10.84 to 19.51), and mean length grew from 4.54 mm (*V. labrusca*) to 6.56 mm (“Izabella”) meaning an increase of 44% (see Table 3).

From the morphological parameters (e.g. roundness and aspect ratio) it can be seen that the shape of *Vitis vinifera* L. varieties are further from a circle than that of *Vitis sylvestris* C. C. Gmel., consequently seeds became elongated due to domestication, which had been confirmed earlier by several researchers (Facsar 2000; Gyulai 2009). The same phenomenon can be observed in the case of *Vitis labrusca* L. (roundness: 0.67) whose cultivated form, called “Izabella”, has a lower roundness value (amounting to 0.58). However, the difference is not as significant as between wild grape (roundness: 0.70) and some of its cultivated varieties (e.g. “Bajor kék”: 0.45), which may indicate the length of time spent in cultivation. This is also verified by “aspect ratio”, which is the ratio of length and width. In the case of longer seeds this value is higher amounting to 1.38 for *Vitis labrusca* L. and 1.61 for “Izabella”. Similarly, “elongation” shows how elongated the seeds are through the ratio of area and perimeter. This value is 2.58 in the case of *Vitis labrusca* L. and amounts to 3.22 for “Izabella”. The variety called “Concord” also originates from the species of *V. labrusca*, and similar trends can be witnessed here too (Lőrincz 2009).

With the help of dimensional and morphological parameters it can be verified that the size of the seeds increased, and they also became elongated and thinner during the artificial cultivation (domestication) of vine, which is also confirmed by *Vitis labrusca* L. and varieties developed (or bred) from it (see Table 3.). This is a well-known and important statement in the case of wine grape.

However, regarding the morphological traits there are some exceptional wine grape varieties (e.g. “Kék őkőrszem”, ”Kéknyelű” and ”Ál-kék hajnos”), where the above mentioned statement is not true, as they have similar ratios as wild grape; but if we also include dimensional parameters in the analysis it becomes unambiguous that they have round but much larger seeds.

Some changes also can be observed when analyzing colour related indices, although they have not been found in the case of wild grape and wine grape. These kinds of analyses are not mentioned even in literature.

Histogram based differentiation – analysed in the case of wild grape and wine grape – also has been carried out for *Vitis labrusca* L. and “Izabella”. The method – which is proven in the case of domesticated plants and their progenitors (Rovner and Gyulai 2007) – lead to the expected results, since the trend lines of the wild species had 3 intersection points, and only 1 intersection point was observed when analysing the domesticated variety. Unfortunately, only a small sample size was available during the investigation of these two items, therefore far-reaching conclusions
cannot be drawn from the results. It is advised to repeat the analysis by using larger sample sizes (meaning at least 150-200 seeds per item).

3.4.6. Differentiation of direct producing grape varieties from other varieties based on the morphometric traits of the seeds

Direct producing varieties could be clearly separated from wild grape and the other *Vitis* species already during the formerly presented discriminant analysis, but they significantly overlapped with wine grape varieties.

According to the further analyses, it seems that from the studied various morphological signs mean hue value (related to the colour of the grape seed) could be suitable for distinguishing the varieties of invasive grape species and the taxons of *Vitis vinifera* L. This parameter includes all the three colour channels, and its values are calculated based on algorithms.

By taking this value into account together with other morphological factors, more accurate results can be achieved. If visualized together with “equivalent diameter” direct producing grape varieties can be clearly separated from the other items in the parameter space. With this method *Vitis sylvestris* C. C. Gmel. items and the other *Vitis* species (*Vitis amurensis, V. labrusca, V. riparia, V. rupestris*) also get separated from the cluster of cultivated *V. vinifera* L. varieties (see Figure 7).

This leads to the conclusion that the colour of the seeds encodes species related properties, and that it is an important factor during the differentiation of direct producing varieties.
Figure 7: Differentiation of direct producing varieties in a double component sample space (direct producing varieties: red, Vitis sylvestris; green, Vitis vinifera varieties: blue, Vitis species: yellow) (own figure, Microsoft Excel 2016)

3.5. New scientific results

1. I collected bunches of grapes from Hungarian variety collections, from which I removed the seeds, and developed a collection that is unique in Hungary and also in Europe. This could be a solid basis for further research, morphological seed analyses and other works (e.g. developing a DNA library or performing analytical tests) besides the preservation of the seeds as genetic resources. Altogether 172 items have been stored including 122 grape varieties and 16 different Vitis species.

2. I elaborated a methodology for the fast and objective morphometric analysis of grape seeds by using easily accessible devices from today’s technology. The method enables for the recording and the storage of the data of recent grape seeds and also those of archaeological origin.

3. I developed an image database of the seeds and the seed groups of the varieties and the species by using the same camera settings. Altogether 82 images have been taken by using 42 items, and they also have been saved in a format prepared for measuring.

4. Dimensional and morphological parameters (56 pieces) – obtained after processing the grape seeds with the presented methodology and tools – have been defined in the case of 6734 seeds, this way providing a basis for further analyses. I summarized the minimum,
maximum and mean values of some of the analysed parameters belonging to the recorded items. This way I could point out the diversity of the *Vitis* genus again.

5. Based on the database I selected those morphometric parameters which are suitable for the differentiation of wild grape and wine grape, and I also could distinguish direct producing grape varieties from *Vitis vinifera* L. varieties by using my method. This enabled for the identification of *Vitis vinifera* L. varieties that can be found in abandoned vineyards, which is the basis for their further analysis. Wine grape, wild grape and the investigated *Vitis* species (*Vitis amurensis*, *V. labrusca*, *V. riparia* and *V. rupestris*) could be clearly differentiated during the discriminant analysis, which could be important for the protection of *Vitis sylvestris* C. C. Gmel.

6. My measurements verified the domestication related morphologic changes of the grape seed in the case of *Vitis vinifera* L. and *Vitis labrusca* L.

7. The species of *Vitis vinifera* L. and *V. sylvestris* C. C. Gmel. cannot be differentiated based on the unimodal or multimodal characteristics of the distribution histograms developed from the values of the different measured parameters. This method used for the identification of the wild and the domesticated characteristics of unknown plant seed populations was not successful in the case of grapes; leading to the conclusion that wine grape has a unique genetic diversity among the cultivated plants, which is also confirmed by its more thousands of varieties.
4. Conclusions and recommendations

During my research work I developed a grape seed collection including the seeds of old Hungarian varieties and other *Vitis* taxons.

Literature review revealed that digital methods and morphometric analyses are widely used for the differentiation of species and varieties. By applying digital image processing I developed a method suitable for taking high quality images of large size samples. After objective computerized processing these images can be used for the comparative analysis of *Vitis* seeds. Due to the continuously expanding database measured parameters can be compared to the stored data, which is useful especially in the case of archaeobotanical data processing.

In the case of wild grape and wine grape the fast measuring of morphometric parameters makes identification easier, and also facilitates the selection of seeds from a larger archaeological sample. The data set compiled of the parameters can be useful during the archaeological analysis of grape seeds, their population level classification and the differentiation and identification of the different varieties. During data assessment the seeds of wild grape and wine grape could be unequivocally distinguished based on the morphometric analysis, and parameters necessary for the identification of direct producing grape varieties also have been defined.

Based on the method it can be stated that dimensional and morphological parameters alone are not suitable for the differentiation of wine grape varieties, therefore seed colour and structure related values also need to be evaluated. By using functions calculated from the parameters with the help of a statistical software (IBM SPSS v.25), *Vitis sylvestris* C. C. Gmel. and invasive *Vitis* taxons spreading on its natural habitats can be also differentiated. This may provide a great opportunity for the protection of the disappearing wild grape, but hybridized populations also need to be included in the analysis. I verified that the analysis of the seeds enables for species level identification, although further research is needed for the elaboration of methods guaranteeing 100% differentiation.

Grapes have plant parts of diverse morphology, which is also true for seeds showing a great variability even within a certain variety. It can be stated that wine grape reached an extreme level of diversity in the Carpathian Basin, which can be also witnessed in the form diversity of the seeds. Regarding the differentiation of the groups of varieties and the species we still have some ongoing investigations. Since the size of the seeds had not been considered during the selection of grapes, and they are not used for propagation either, it seems that mostly genetic factors influence their traits. Although their form and shape are regarded to be species or variety specific, it would be also important to investigate the role of environmental impacts in the development of their traits.
In the case of conducting further analyses it could be explored how the characteristics of the parents are manifested in the new varieties developed through breeding and crossings. This would also prove the suitability of the method for defining relationships.

As a result of my investigations, I defined the minimum, maximum and mean values of the most important parameters of the measured items. These can be used for reaching more accurate results during identification. Based on the registered variables, the number of possible cultivars can be decreased, and some varieties can be excluded when analyzing an unknown archaeological (or recent) sample. By complementing historical data, even varieties can be more or less identified in some cases. Since my objective was to elaborate a method for the fast identification of large size samples, some important variables (hilum, ventral furrow or beak length) were not included in the analyses, although they may carry species or variety specific information. These variables cannot be considered by my method, since the software is not suitable for their accurate recognition. However, in the case of small size samples their analysis can be performed manually, making the results of digital data processing more accurate.

I advise to collect and to process the seeds of the most important varieties stored in the different Hungarian gene banks for the expansion of the digital image and morphometric database. Further research is needed to explore and make use of the possibilities provided by the morphometric analysis of seeds. Hungarian research institutes of viticulture and oenology and the National Centre for Biodiversity and Gene Conservation located in Tápiószele could be possible partners in this process.

We started to integrate the more than 6000 measured image data into an artificial intelligence (AI) based decision making system. This AI system is already suitable for the recording of the visual and the numeric data of large sample size archaeological findings, the classification of the items, and variety or variety group level identification. By expanding the database, the method may allow for the development of an automated identification system. This could be carried out by using artificial neural networks (ANN), which can be applied in numerous ways, for example for performing other image processing tasks. They will be able to recognize unknown samples through measuring or learning (training), but not by programming. This could be a great help for those users not experienced in programming. Since the database can be continuously expanded, classification is to become more and more accurate over time.
5. Literature


6. Scientific publications covering the topic of the thesis

Hungarian books or book chapters:


Proofread foreign language journal articles:


Proofread Hungarian journal articles:


Full abstracts published in conference publications (Hungarian):


Abstracts published in conference publications (foreign language):


Abstracts published in conference publications (Hungarian):


