taxonomy

MDPI

Review

When Genetics Meets Ecology: Genomics and Taxonomy of *Vitis* Species and Cultivars

José Luis Rodríguez Lorenzo [1] and Emilio Cervantes [2,*]

[1] Plant Developmental Genetics, Institute of Biophysics v.v.i, Academy of Sciences of the Czech Republic, Královopolská 135, 612 65 Brno, Czech Republic; rodriguez@ibp.cz
[2] Instituto de Recursos Naturales y Agrobiología de Salamanca (IRNASA), Consejo Superior de Investigaciones Científicas (CSIC), Cordel de Merinas, 40, 37008 Salamanca, Spain
* Correspondence: emilio.cervantes@irnasa.csic.es

Abstract

Knowledge of the biology of the genus *Vitis* has undergone a profound transformation, evolving from the morphological descriptions of classical ampelography to the high-resolution analyses enabled by modern phylogenomics. This review explores the "Paradox of the Vine"—the remarkable phenotypic plasticity that historically complicated botanical nomenclature—and examines how genomic tools have helped resolve many of these long-standing taxonomic challenges. We trace the development of grapevine genomics from the first near-homozygous reference genome (PN40024) to the current era of telomere-to-telomere (T2T) assemblies and phased diploid genomes. Attention is given to the genomic "dark matter" represented by transposable elements and structural variation, which contribute substantially to varietal identity and species-specific adaptation to changing environmental conditions. Advances in bioinformatic methodologies, including pangenome graph construction and machine learning-based variant detection, now enable clonal discrimination and complex parentage analysis with unprecedented precision. The definition of genuine wild grapevines (*Vitis vinifera* subsp. *sylvestris*) remains a critical issue in studies of grapevine evolution, domestication, and genome structure. The traditional concept of wild populations free from introgression by cultivated grapevines has been increasingly challenged by ecological observations and molecular evidence. Distinguishing truly wild populations from feral lineages is therefore essential for reconstructing the history of grapevine domestication and understanding patterns of gene flow between cultivated and wild compartments. Future progress in *Vitis* systematics will depend on the integration of genomic, ecological, and morphometric approaches. We propose that the next generation of grapevine taxonomy will combine the historical insights of ampelography with high-throughput phenotyping and comprehensive pangenomic resources, leading to a predictive and evolutionarily informed framework for the classification of *Vitis* species and cultivars.

Keywords: ampelography; assembly; clonal identification; curvature; ferals; morphometry; pangenome; seeds; structural variation

Check for updates

Academic Editor: Mathias Harzhauser

Received: 4 June 2026
Revised: 23 June 2026
Accepted: 24 June 2026
Published: 27 June 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. Evolutionary Complexities and Traditional Taxonomy in the Family Vitaceae

The family Vitaceae Juss. comprises sixteen genera and approximately 950 species of woody and herbaceous climbing vines distributed predominantly throughout tropical and

subtropical ecosystems. Together with Leeaceae Dumort., it constitutes the order *Vitales*, one of the earliest-diverging lineages within the Rosids, whose closest phylogenetic relatives remain a matter of debate [1–3]. The evolutionary history of the family extends from its stem lineage in the Late Cretaceous (ca. 90 Ma) to the diversification of extant crown groups during the Oligocene (ca. 30 Ma) [4,5]. The family Vitaceae is currently divided into five tribes: (I) Ampelopsideae (comprising *Ampelopsis* Michx., *Nekemias* Raf., *Rhoicissus* Planch., and *Clematicissus* Planch.); (II) Cisseae, represented by *Cissus* L.; (III) Cayratieae, including *Cayratia* Juss. ex Guill, *Causonis* Raf., *Acareosperma* Gagnep., *Afrocayratia* J.Wen, L.M.Lu, Rabarij. & Z.D.Chen, *Cyphostemma* (Planch.) Alston, *Pseudocayratia* J. Wen, L.M.Lu & Z.D.Chen and *Tetrastigma* Planch); (IV) Parthenocisseae, represented by *Parthenocissus* Planch.; and (V) Viteae. The latter tribe comprises three closely related genera: *Ampelocissus* Planch., *Pterisanthes* Blume, and *Vitis* L., the economically and evolutionarily most significant genus within the family [2–4].

While the deep evolutionary history of the family is anchored by macrofossil discoveries from the Late Cretaceous Deccan Cherts of India (ca. 66 Ma) [6] and Paleocene deposits in the Neotropics [1,5], resolving lower-level phylogenetic relationships remains challenging. Within the tribe Viteae, the genus *Vitis* comprises approximately 80 species currently recognized by the expert consensus curated in Kew's Plants of the World Online (POWO), the majority of which are dioecious climbing plants [7].

Large-scale macroecological analyses have demonstrated that dioecy is significantly associated with the climbing habit, woodiness, and the production of fleshy fruits across angiosperms [8]. This pattern suggests that the reproductive biology of grapevines is closely linked to their ecological role as forest lianas. Consequently, the transition from dioecious wild populations to predominantly hermaphroditic cultivated forms represents one of the most profound evolutionary shifts associated with grapevine domestication. This transition not only altered reproductive biology but also reshaped patterns of genetic diversity, gene flow, and human-mediated selection within the genus.

1.2. Biogeography and Phylogenetics of the Genus Vitis

The genus *Vitis* is divided into two subgenera that differ in both chromosome number and morphology: subgenus *Muscadinia* (n = 20), comprising *Vitis rotundifolia* Michx. and *Vitis popenoei* J.L. Fennell, and subgenus *Vitis* (n = 19), which includes most extant species [9–12] (Figure 1). The distinction between these subgenera is supported by cytogenetic, morphological, and molecular evidence and reflects an early evolutionary divergence within the genus. Despite their shared ancestry, species belonging to the two lineages differ markedly in chromosome structure, reproductive compatibility, and patterns of geographic distribution, making them important models for investigating speciation and genome evolution in grapevines.

The geographic distribution of wild *Vitis* species is closely associated with their underlying genetic structure. Analyses based on multilocus nuclear markers and plastid DNA sequences consistently support the monophyly of the genus and resolve subgenus *Vitis* into several geographically distinct evolutionary lineages [11,12]. Natural hybrid zones, particularly widespread in North America, were recognized early in viticultural history through the observation of intermediate morphological and agronomic characteristics [13].

Modern phylogenetic studies have clarified these relationships, identifying several well-defined North American clades, including the Aestivales group (e.g., *Vitis aestivalis* L. and *V. riparia* Michx.) and the Labruscae group (e.g., *V. labrusca* L. and *V. mustangensis* Buckley), which are clearly differentiated from a strongly supported monophyletic Asian lineage [11]. These biogeographic patterns indicate a long history of regional diversification

and provide an essential framework for understanding the evolution, adaptation, and domestication of grapevines.

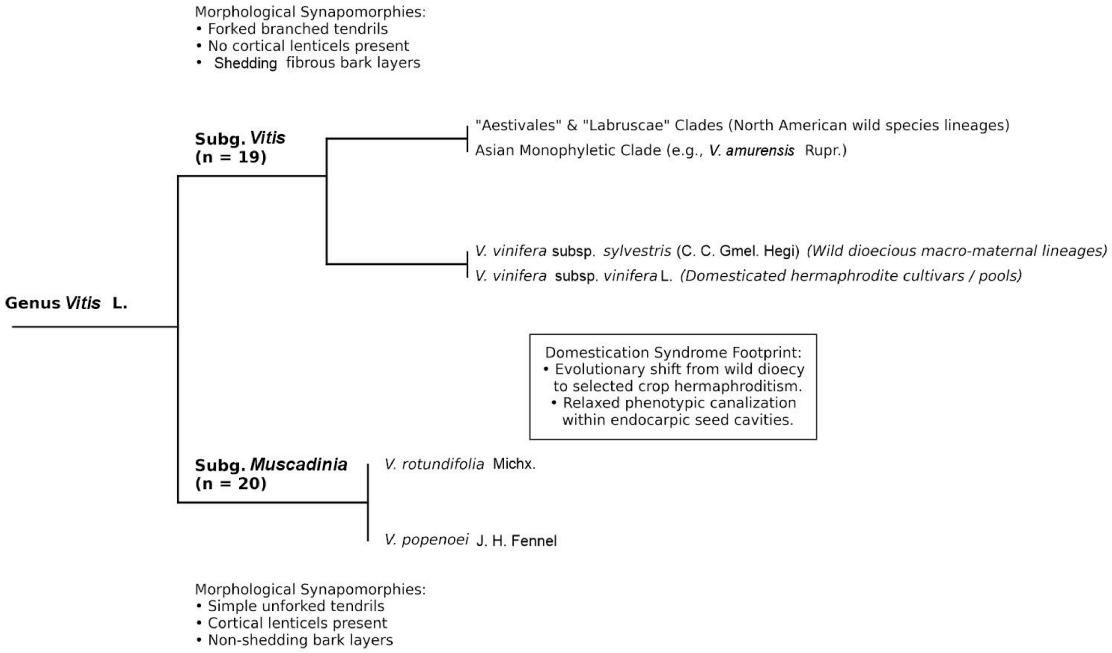

Figure 1. Phylogenetic relationships and diagnostic morphological characteristics of the genus *Vitis* L. The schematic phylogeny illustrates the divergence between the two recognized subgenera, *Vitis* (n = 19) and *Muscadinia* (n = 20). Diagnostic morphological features, including tendril branching patterns, the presence of cortical lenticels, and bark exfoliation characteristics, are indicated for each lineage. Terminal branches represent the principal evolutionary groups within the genus, including the North American clades, the Asian monophyletic clade, *Vitis vinifera* subsp. *sylvestris* (Willd.) Hegi, and cultivated grapevines (*V. vinifera* subsp. *vinifera*). The diagram highlights the phylogenetic context of grapevine domestication and the emergence of traits associated with the domestication syndrome.

The final major evolutionary lineage within the genus comprises the Euro-Mediterranean grapevines: the wild *Vitis vinifera* subsp. *sylvestris* (Willd.) Hegi and the domesticated *V. vinifera* subsp. *vinifera* (syn. subsp. *sativa*). The persistent phylogenetic clustering of cultivated varieties with geographically proximate *sylvestris* populations reflects their shared ancestry and provides important evidence for the origins of grapevine domestication [14,15]. At the same time, this close genetic relationship complicates taxonomic and conservation efforts by obscuring the distinction between genuinely wild populations and feral individuals derived from escaped cultivated vines. Such feral populations may introgress into native gene pools and thereby distort assessments of wild genetic diversity and population structure.

Molecular analyses based on microsatellite (SSR) markers have revealed substantial genetic affinities between wild and cultivated grapevines across the Caucasus and Central Asia, including populations from Georgia, Uzbekistan, Tajikistan, and Kyrgyzstan [16]. Similarly, high-density SNP genotyping using the Vitis18k array has identified three principal *V. vinifera* subsp. *sylvestris* genetic groups, with North African–Iberian populations occupying an intermediate position between the Georgian lineage and populations from Western and Central Europe [17,18]. These findings support a complex demographic history characterized by migration, introgression, and recurrent gene flow between cultivated and wild grapevine populations throughout the Euro-Mediterranean region.

1.3. From Ampelography to Genomics

Historically, diagnostic traits such as tendril morphology (simple versus bifurcate), the presence of cortical lenticels, and variation in leaf architecture formed the foundation of ampelography, the botanical discipline devoted to the identification, description, and classification of grapevine cultivars [19–21]. The field was formalized through the pioneering work of authors such as Simón de Rojas Clemente y Rubio [19], Pierre Galet [20], and Viala and Vermorel [21]. In recent decades, traditional ampelographic methods have been substantially enhanced through the integration of digital image analysis, geometric morphometrics, multivariate shape quantification, and quantitative trait locus (QTL) mapping of leaf morphology [22–24].

Despite these advances, morphological descriptors remain subject to important limitations, particularly when applied to cultivated grapevines. Leaf and seed morphology in *Vitis vinifera* exhibit considerable phenotypic plasticity in response to environmental conditions, vine age, and edaphic factors [23], often obscuring the boundaries between closely related cultivars. Moreover, traditional taxonomic approaches face three major challenges:

- Synonymy and Homonymy. Genetically identical cultivars may be maintained under different regional names, as exemplified by the well-documented synonymy of Zinfandel, Primitivo, and Crljenak Kaštelanski. Conversely, distinct genetic lineages may share the same cultivar designation, generating complex homonymic groups such as those associated with Malvasia or Lambrusco
- Clonal Variation. Somatic mutations accumulated during vegetative propagation can modify traits of agronomic and oenological importance, including berry colour, fruit composition, growth habit, and disease resistance. Such variation is frequently undetectable through conventional morphological examination alone.
- Reticulate Evolution. Extensive hybridization and introgression among species and cultivars have generated complex reticulate evolutionary patterns, particularly in regions where multiple taxa occur in sympatry. These processes often obscure parental relationships and challenge the assumptions of strictly bifurcating phylogenetic models [25].

The emergence of molecular markers and, more recently, whole-genome sequencing has transformed grapevine taxonomy by providing objective tools for cultivar identification, parentage reconstruction, clonal discrimination, and the analysis of evolutionary relationships. As a result, modern grapevine systematics increasingly integrates classical ampelographic observations with genomic evidence, bridging the gap between phenotype and genotype.

In this review, we examine how contemporary grapevine genomics and pangenomics address longstanding taxonomic and evolutionary challenges, providing a framework for a predictive and high-resolution classification of the genus *Vitis*. We begin by establishing the genomic foundation of modern grapevine research through the near-homozygous PN40024 reference genome and trace its evolution toward complete telomere-to-telomere (T2T) chromosome assemblies [26]. Moving beyond single-reference genomes, we explore how structural variations (SVs), copy-number variations (CNVs), and pangenome graph representations have revealed previously hidden dimensions of grapevine diversity, shedding light on the processes of domestication and the complex genetic relationships among wild, cultivated, and feral lineages.

Finally, we discuss the implications of these advances for contemporary viticultural taxonomy and systematics. We argue that the integration of phased genome assemblies, pangenomic resources, high-throughput morphometric analyses, and ecological data will enable a new generation of classification frameworks that more accurately reflect the evolutionary history, genetic diversity, and adaptive potential of *Vitis* species and cultivars.

2. The Molecular Shift: From Ampelography to Chloroplast Genomics

2.1. Classification of Vitis Cultivars in the Early Molecular Era

To overcome limitations imposed by morphological plasticity, the introduction of molecular markers at the turn of the 21st century revolutionized viticultural systematics. Nuclear tools—initially multi-allelic microsatellites (SSRs) and later high-density SNP arrays—provided the first robust, empirical frameworks to resolve complex pedigrees, reconstruct reticulate networks, and identify core ancestral founders at the root of extensive varietal lineages [27]. Landmark parentage studies successfully elucidated the origin of renowned cultivars such as the origin of Cabernet Sauvignon from Cabernet Franc and Sauvignon Blanc [28]. The scalability of this approach was subsequently demonstrated through large-scale pedigree reconstructions, which revealed that a relatively small number of founder cultivars contributed disproportionately to the ancestry of thousands of contemporary European varieties [29,30]. Notably, in the Iberian Peninsula, the ancient, functionally female cultivar Hebén has been identified as a major founder genotype, occupying a central position in the pedigree network of numerous traditional Spanish cultivars and contributing substantially to the genetic structure of western European grapevine diversity [31–33] (Figure 2).

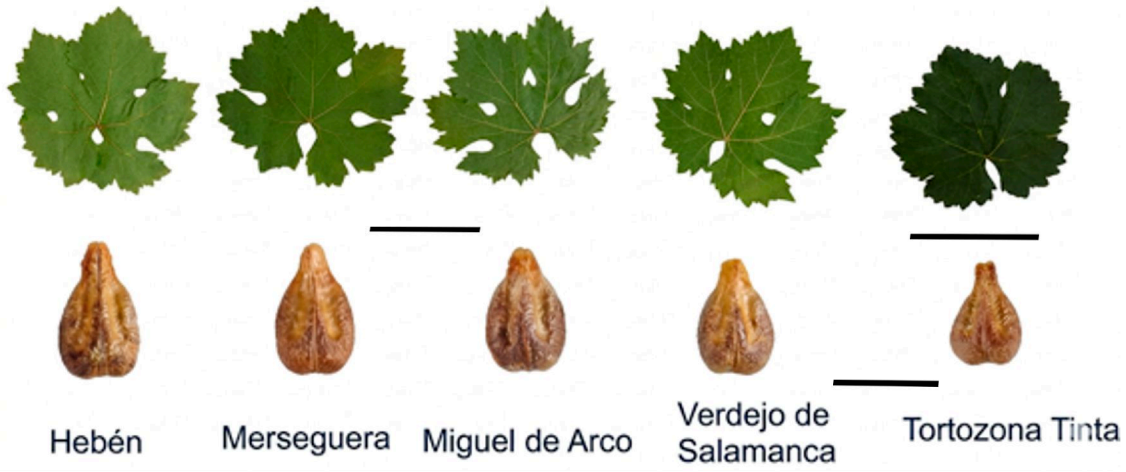

Figure 2. Leaves and seeds of *Vitis vinifera* cv. Hebén, three cultivars from its progeny (Merseguera, Miguel de Arco, and Verdejo de Salamanca), and the unrelated cultivar Tortozona Tinta [31]. All samples were collected at IMIDRA during the 2024 season. The leaves and seeds of Hebén and its progeny share a higher morphological resemblance to each other than to those of the unrelated cultivar, Tortozona Tinta. Scale bars represent 10 cm for the leaves of Hebén, Merseguera, Miguel de Arco, and Verdejo de Salamanca, and 5 cm for Tortozona Tinta. The bottom scale bar represents 5 mm for all seeds.

Prior to the advent of molecular approaches and the definitive genetic resolution of cultivar relationships, early classifications of the extensive intraspecific diversity within *Vitis vinifera* were based primarily on eco-geographical and morphological frameworks. Among the pioneers of this approach was Nikolai Vavilov, whose investigations into global crop biodiversity led to the hypothesis that regions exhibiting the greatest genetic diversity of a crop and its sympatric wild relatives corresponded to its primary center of domestication [34]. Applying this concept to the grapevine, Vavilov proposed that its cultivation originated in Inner Asia (Center III) and Asia Minor (Center IV) [34]. Building upon this framework, Alexander Negrul integrated extensive ampelographic, physiological, and morphometric evidence to partition *V. vinifera* into three major eco-geographical groups, termed *proles* [35]:

- Proles occidentalis: Comprising wine grapes from Western Europe (e.g., Cabernet, Pinot, Riesling, Traminer), typically characterized by small bunches and small berries with high acidity.
- Proles pontica: Originating from the Black Sea and Balkan regions (e.g., Clairette, Furmint, Kisi), exhibiting intermediate physiological and morphological features.
- Proles orientalis: Primarily table grapes from the Near and Far East (e.g., Sultanina, Muscat), distinguished by large, fleshy berries with firm pulp and thin skins. Negrul further divided this group into the ancient sub-group *Caspica* (historically utilized for early vinification) and *Antasiatica* (selected for table use and raisins, displaying historical trade links to the Western Mediterranean).

2.2. Plastid Phylogeography and Feralization Dynamics

While nuclear markers capture patterns of biparental inheritance and recombination, the reconstruction of deep maternal lineages and broad-scale domestication pathways relied primarily on analyses of the chloroplast genome (cpDNA). Owing to its largely non-recombining nature, conserved structural organization, and predominantly maternal inheritance in the Vitaceae, the chloroplast genome provides a valuable record of evolutionary history and maternal ancestry [4,12,36].

Early restriction fragment length polymorphism (RFLP) studies, followed by analyses of chloroplast microsatellites (cpSSRs) and chloroplast sequence variation, revealed strong phylogeographic structuring within *Vitis vinifera* and supported a model involving a primary domestication center in the Near East, accompanied by subsequent regional introgression and diversification within the Western Mediterranean basin [36]. Across these regions, the distribution of the four principal chloroplast haplotypes (chlorotypes A, B, C, and D) exhibits distinct biogeographical patterns. Chlorotypes C and D predominate in eastern germplasm groups, particularly those associated with *proles orientalis*, whereas chlorotypes A and B are more frequently observed in western *V. vinifera* subsp. *sylvestris* populations and traditional cultivars from the Iberian Peninsula and Western Europe, corresponding broadly to the *proles occidentalis* group [34–36]. These phylogeographic patterns provide molecular support for the eco-geographical differentiation originally proposed by Vavilov and later formalized by Negrul [34,35].

Despite their phylogeographic utility, chloroplast haplotypes also introduce specific challenges for interpreting domestication history, feralization processes, and asymmetric gene flow. Wild *Vitis vinifera* subsp. *sylvestris* populations are dioecious, comprising separate male and female individuals, whereas cultivated grapevines are predominantly hermaphroditic [37]. Consequently, the distinction between genuinely wild populations and feral descendants of cultivated vines cannot always be resolved through chloroplast markers alone. Recent high-throughput studies integrating cpSSRs, cpSNPs, and nuclear genomic data have demonstrated ongoing bidirectional gene flow between wild and cultivated compartments, mediated by pollen transfer from wild male vines to cultivated vineyards and by the establishment of feral individuals derived from cultivated maternal lineages within natural riparian habitats [38,39].

This continuous genetic exchange has generated complex patterns of admixture across Mediterranean and Central Asian grapevine populations, often blurring the boundaries between wild, cultivated, and feral gene pools. Consequently, the identification of authentic relict *V. vinifera* subsp. *sylvestris* populations increasingly relies on integrated genomic approaches capable of jointly analyzing cytoplasmic and nuclear variation. Such approaches, particularly those based on complete, gapless genome assemblies, offer the resolution required to disentangle ancient ancestry from more recent introgression and feralization events [14,15].

3. PN40024 and the Near-Homozygous Model: Establishing a Genomic Baseline

Cultivated *Vitis vinifera* exhibits high levels of genomic heterozygosity relative to many plant reference genomes, reflecting its long evolutionary history of outcrossing and the widespread use of vegetative propagation in viticulture [14,40]. Recurrent clonal reproduction has preserved numerous heterozygous allelic combinations that would otherwise be reshuffled through sexual reproduction, contributing to the complex genomic architecture characteristic of domesticated grapevine cultivars [41]. This extensive heterozygosity represented a major challenge for early genome assembly efforts. Both cultivated and wild grapevines harbour substantial sequence divergence between homologous chromosomes, including structural variants (SVs), insertions and deletions (indels), and other forms of haplotypic variation. Whereas such diversity is maintained in wild populations through obligate outcrossing associated with dioecy, centuries of clonal propagation have preserved these divergent haplotypes within cultivated lineages. Consequently, early de novo assembly strategies based primarily on short-read sequencing data often struggled to accurately reconstruct highly heterozygous genomic regions, frequently collapsing divergent haplotypes into consensus sequences or generating fragmented assemblies.

3.1. The Selection and Construction of the Model

To overcome the challenges that high heterozygosity posed for early genome assembly, the French–Italian Public Consortium for Grapevine Genome Characterization sought to develop a reference genotype with a substantially reduced level of heterozygosity. The outcome of this strategy was PN40024, a specialized line generated specifically for genomic research rather than for agronomic use. PN40024 traces its origin to Helfensteiner, a dark-skinned cultivar bred in 1931 from a controlled cross between Schiava Grossa (Trollinger) and Pinot Noir [42]. Researchers at the Institut National de la Recherche Agronomique (INRA, Colmar, France) subjected this lineage to nine successive generations of self-fertilization (S9) [43]. This extensive inbreeding program dramatically reduced heterozygosity, producing a near-homozygous genotype in which approximately 93% of loci were estimated to be fixed [43]. As a result, much of the allelic and structural complexity that characterizes the grapevine genome was minimized, providing a tractable template for the development of the first reference assembly.

3.2. A Landmark in Plant Phylogenomics

Released in 2007, the PN40024 draft genome represented a major milestone in plant genomics. It was the first fruit crop genome to be sequenced and assembled, and only the fourth angiosperm genome available to the scientific community [43]. In parallel, an independent whole-genome shotgun assembly was generated for the highly heterozygous cultivar Pinot Noir (clone ENTAV 115), providing an important point of comparison for assessing the consequences of natural heterozygosity on genome assembly and structural variation [44]. More recently, advances in long-read sequencing technologies have enabled the reconstruction of gapless reference assemblies for grapevine cultivars, including Pinot Noir, substantially improving the resolution of complex genomic regions and haplotypic structure [26].

The establishment of the standardized PN40024 reference genome provided three immediate and enduring foundations for grapevine genomics:

- A Standardized Genomic Reference Framework: For the first time, grapevine researchers possessed a common genomic coordinate system that enabled the precise localization of genes, quantitative trait loci (QTLs), and molecular markers across the 19 grapevine chromosomes ($2n = 38$; $n = 19$). This framework facilitated the integra-

tion and comparison of genetic information generated from cultivated varieties, wild accessions, and interspecific hybrids.

- A Platform for Genome-Wide Comparative Analyses: The PN40024 assembly accelerated the transition from sparse microsatellite (SSR) markers to high-density single-nucleotide polymorphism (SNP) datasets, enabling large-scale studies of pedigree reconstruction, population structure, domestication history, and genome-wide patterns of selection and structural variation [45].
- Insights into Angiosperm Genome Evolution: Comparative analysis of the grapevine genome provided important evidence for the ancestral paleohexaploidy event (γ), a whole-genome triplication shared by core eudicots. Because the grapevine genome has undergone relatively limited lineage-specific polyploidization since this event, *Vitis vinifera* emerged as a valuable reference for reconstructing the evolutionary history and genomic architecture of flowering plants [43].

Despite its foundational role as the reference genotype for grapevine genomics, relatively little information is available regarding the physiological characteristics of PN40024 under field conditions. This limited phenotypic characterization reflects its development as a research resource rather than as a cultivar intended for agricultural production. Nevertheless, the extensive inbreeding required to generate PN40024 substantially reduced genome-wide heterozygosity, producing a genotype that exhibits characteristics commonly associated with inbreeding depression. Compared with highly heterozygous cultivated grapevine varieties such as Pinot Noir, PN40024 generally displays reduced vegetative vigour, lower reproductive performance, and diminished overall fitness, highlighting the biological costs associated with prolonged self-fertilization in an otherwise highly outcrossing species.

4. Growth and Development of the Model

Although the original PN40024 reference genome provided the essential foundation for modern grapevine genomics, it also highlighted important limitations inherent to first-generation reference assemblies. By design, PN40024 represented a highly homozygous genotype and was assembled as a largely linear reference sequence, facilitating genome reconstruction and annotation but only partially capturing the extensive genomic diversity characteristic of the genus *Vitis*. Consequently, many forms of genetic variation—including structural variants (SVs), hemizygous regions, and highly divergent haplotypes—were underrepresented or absent from the initial assembly. While this simplification was necessary to overcome the technical constraints of early sequencing and assembly technologies, it also limited the ability of the reference genome to fully reflect the genetic complexity underlying phenotypic diversity, environmental adaptation, and evolutionary processes in grapevine populations.

As genomic technologies advanced, these limitations became increasingly apparent, motivating the development of improved assemblies, haplotype-resolved genomes, and ultimately pangenomic frameworks designed to capture the full spectrum of variation present within *Vitis vinifera* and related species.

While PN40024 provided an indispensable reference for the conserved genomic architecture of grapevine and other core eudicots, it captured only a portion of the extensive structural and allelic diversity present within the species. Highly divergent haplotypes, structural variants (SVs), hemizygous regions, and cultivar-specific insertions, deletions, and transpositions were often difficult to resolve using the short-read sequencing technologies available at the time. As a result, many of the genomic features underlying phenotypic variation, environmental adaptation, and cultivar differentiation remained incompletely characterized.

To address these limitations, the original PN40024 assembly was progressively refined through the integration of high-density genetic maps and physical mapping resources, culminating in the release of the 12X.v2 assembly [46]. This updated version substantially improved assembly continuity and chromosome-scale organization, providing a more robust framework for genome annotation and comparative analyses. Nevertheless, the full resolution of the grapevine genome structure became possible only with the advent of long-read sequencing technologies. These advances ultimately led to the publication of the first telomere-to-telomere (T2T) assembly of *Vitis vinifera*, representing a major step toward the complete characterization of grapevine genome architecture [26].

4.1. Resolving the Genome: Telomere-to-Telomere (T2T) Assemblies and Phased Haplotypes

Because *Vitis* species and most cultivated grapevine varieties are highly heterozygous and predominantly outcrossing, a single linear reference genome can only partially represent the full extent of genomic variation present within an individual. Advances in long-read sequencing (LRS) technologies, coupled with chromosome-conformation and phasing methodologies, have enabled the reconstruction of haplotype-resolved genomes, in which the two homologous chromosome sets inherited from each parent are assembled separately. This approach preserves allelic diversity and structural variation that are frequently obscured in consensus assemblies.

The current benchmark in genome assembly is the telomere-to-telomere (T2T) framework, which aims to reconstruct complete chromosome sequences from one telomere to the other without unresolved gaps. Such assemblies provide unprecedented access to previously inaccessible genomic regions, including centromeres, telomeres, large repetitive elements, and complex structural rearrangements. In grapevine, these high-resolution assemblies have expanded our capacity to investigate genome evolution, structural variation, haplotype diversity, and the genetic mechanisms underlying adaptation, domestication, and cultivar differentiation.

The emergence of long-read sequencing (LRS) technologies—driven primarily by Pacific Biosciences (PacBio) High-Fidelity (HiFi) circular consensus sequencing and Oxford Nanopore Technologies (ONT) ultra-long reads—has transformed grapevine genomics by enabling the characterization of genomic regions that were difficult to resolve using short-read approaches. PacBio HiFi sequencing combines long read lengths with high per-base accuracy, whereas ONT platforms can generate ultra-long reads capable of traversing highly repetitive genomic regions and complex structural rearrangements. In *Vitis*, these technologies routinely produce reads exceeding 15–20 kb and, in some cases, substantially longer sequences. Such extended read lengths facilitate the assembly of transposable elements (TEs), tandemly duplicated gene clusters, and other repetitive regions that frequently fragment short-read assemblies. Consequently, LRS has greatly improved the completeness, contiguity, and structural accuracy of grapevine genome assemblies, providing a more comprehensive representation of genome organization and genetic variation.

- PN40024.v4 Assembly: Velt et al. [42] released a substantially improved version of the grapevine reference genome, incorporating 35,230 annotated protein-coding genes together with an alternative haplotype assembly. Beyond improving assembly quality and annotation accuracy, this work clarified the genetic origin and pedigree of PN40024, confirming its derivation from the Schiava Grossa × Pinot Noir lineage through the cultivar Helfensteiner [42].
- PN40024 Telomere-to-Telomere (T2T) Reference: Shi et al. [26] reported the first complete T2T assembly of the PN40024 genome, resolving centromeric, telomeric, and other highly repetitive regions that had remained incompletely assembled in previous reference versions. This achievement revealed the extensive contribution of repetitive

DNA to grapevine genome architecture, with repetitive sequences accounting for approximately 67% of the assembled genome, including centromeric repeats, telomeric regions, and ribosomal DNA (rDNA) arrays [26].

The remarkable consistency of the approximately 67% repetitive DNA content observed across the independent telomere-to-telomere (T2T) assemblies of PN40024, PN1, and PN2 [26,41] suggests a highly conserved structural organization of the grapevine genome. This pattern may reflect the long-term stability of repeat-rich, low-recombination genomic compartments, particularly within pericentromeric and other heterochromatic regions. Beyond their structural role, these repetitive elements constitute a substantial fraction of the grapevine genome and are increasingly recognized as important contributors to chromatin organization, genome evolution, and gene regulation.

In earlier reference assemblies, the full extent of transposable element (TE) diversity and organization was difficult to resolve. Because many TEs consist of long, highly similar sequences—often extending over several kilobases— the short reads generated by Sanger sequencing and early Illumina platforms were frequently insufficient to span entire repetitive regions. Consequently, genome assembly algorithms commonly encountered two major challenges:

- Repeat Collapse: Multiple highly similar repeat copies were merged into a single consensus sequence, leading to an underestimation of repeat abundance and reducing assembly accuracy within repetitive regions.
- Assembly Gaps and Fragmentation: When repetitive sequences could not be confidently reconstructed, assemblies frequently contained unresolved regions represented by stretches of ambiguous bases (N's), resulting in fragmented genome maps and incomplete structural information.

The original PN40024 assembly, constrained by the sequencing technologies available at the time, primarily captured the unique and low-copy regions of the genome, particularly protein-coding sequences and their immediate genomic context. Consequently, a substantial proportion of the repetitive genome remained incompletely resolved. This repeat-rich component—sometimes referred to as the genomic "dark matter"—includes transposable elements (TEs), satellite repeats, and other highly repetitive sequences that play important roles in genome structure, evolution, and regulation.

Among these elements, Class I retrotransposons of the Copia superfamily, particularly members of the Ale lineage, constitute a major fraction of the grapevine repeatome [47]. Comparative genomic analyses suggest that these elements underwent extensive amplification during the evolutionary history of *Vitis*, contributing substantially to genome expansion and structural diversification [43,47]. Their insertion activity has been associated with genomic regions involved in metabolism, stress responses, and other adaptive processes, highlighting their potential role in shaping patterns of genetic variation. In addition, horizontal transfer has been proposed as a mechanism contributing to the evolutionary dissemination of certain transposable element lineages across plant genomes, including those present in ancestral *Vitis* lineages [48].

4.2. General Landscape of the Vitis Genome Based on the Model

The grapevine genome is characterized by a relatively compact size combined with a substantial proportion of repetitive DNA. Table 1 summarizes the principal structural features of the PN40024 reference genome and its successive improvements. Specifically, Table 1a illustrates the progression of gene annotation across different assembly versions; Table 1b summarizes the major genomic characteristics of the reference assembly, including genome size, chromosome organization, and repetitive content; and Table 1c presents key gene families associated with seed development and related developmental processes.

Table 1. (**a**) Number of genes in the successive versions of PN40024. (**b**) Characteristics of the genome of PN40024, the model for *Vitis*. (**c**) Key gene families involved in seed development (Based on PN40024.v5 T2T, [26]).

(**a**)

Genome Assembly Version	Curation Year	Total Protein-Coding Genes	Net Gene Variance Relative to Original Assembly	Refs.
12X.v1 (CRIBI/IGGP)	2007/2011	29,971	*Baseline Reference Draft*	[46]
v4 (Integra/INRAE)	2019	31,000	+1029 newly validated open-reading frames	[42]
v5 (T2T-GAPLESS/Grapedia)	2023	39,234	+9263 structural loci resolved via long-reads	[26]

(**b**)

Structural Feature	Physical Specifications & Biological Context	Refs.
Genome Size & Karyotype	The haploid genome spans an estimated 475 to 500 Mb distributed across 19 standard chromosomes (n = 19). While physically small compared to major monocot systems like maize or wheat, its complexity is driven by its highly heterozygous, outbred nature.	[44]
Coding Capacities	Approximately ~30,000 baseline protein-coding loci. Representative genes mapped to distinct chromosomal coordinates include: • *VvGAI1* (DELLA growth-repressor protein family; Chromosome 1) • *VvMybA1/VvMybA2* (Master switch regulatory cluster for anthocyanin fruit pigmentation; Chromosome 2) • *VvTPS* (Terpene Synthases driving primary aroma volatilization profiles; Chromosome 15) • *R Genes* (NBS-LRR structural disease-resistance clusters; Chromosome 16) • *VvSTS* (Stilbene Synthases governing primary phytoalexin defense biosynthesis; Chromosome 18)	[44]
Heterochromatin Distribution	Over 50% of the complete physical DNA landscape consists of repetitive blocks and dense non-coding structures heavily localized around structural centromeric and pericentromeric domains.	[26]
Transposable Elements (TEs)	Dominance of Class I Retrotransposons, specifically Long Terminal Repeat (LTR) architectures. The two most abundant superfamilies are *Copia* and *Gypsy*, which account for most TE-derived genomic mass. These elements are non-randomly integrated, acting as the primary drivers of intergenic region expansions and variable structural indels.	[26]

(**c**)

Gene Family	Total Family Loci Count (v5)	Primary Mechanistic Role in Seed Development	Representative Target Loci Examples	Foundational Literature References
MADS-box	110	Master regulators governing ovule identity, floral organ determination, and maternal integument (seed coat) expansion.	*VvAGL11*, *VvSTK*	Mejía et al. (2011); Grimplet et al. (2014); Shi et al. (2023) [26,49,50]
MYB Factors	300	Fine-tuned transcriptomic orchestration of tannin, proanthocyanidin, and flavonoid pathways inside developing seed coats.	*VvMYBPA1*, *VvMYB5a*	Bogs et al. (2007); Cavallini et al. (2014) [51,52]
NAC Domain	160	Programmed cell wall thickness regulation, secondary lignin deposition, and structural seed coat hardening.	*VvNAC26*, *VvNAC44*	Wang et al. (2013); Grimplet et al. (2012) [49,53]

Table 1. *Cont.*

(c)				
Gene Family	**Total Family Loci Count (v5)**	**Primary Mechanistic Role in Seed Development**	**Representative Target Loci Examples**	**Foundational Literature References**
ARF (Auxin Response)	30	Downstream mediation of endogenous auxin signaling driving cell expansion inside the seed embryo and surrounding berry pericarp.	*VvARF2*, *VvARF4*	Wan et al. (2014) [54]
B3 Domain	80	Transcriptomic control over structural embryo maturation phases and internal storage reserve deposition.	*VvL1L*, *VvABI3*	Ahmad et al. (2019) [55]
TCP Family	25	Coordinated cell proliferation dynamics, growth symmetry preservation, and directional morphological development of the seed.	*VvTCP1*, *VvTCP15*	Jiu et al. (2019); Shi et al. (2023) [26,56]

4.3. Quantifying Diversity and the Inbreeding Landscape

The availability of telomere-to-telomere (T2T) and haplotype-resolved assemblies has provided unprecedented opportunities to examine how different reproductive strategies—including outcrossing, self-fertilization, and clonal propagation—shape genome-wide patterns of variation in grapevine. In particular, the contrast between the naturally highly heterozygous condition of most *Vitis* populations and the near-homozygous state of PN40024 offers a valuable framework for investigating the genomic consequences of inbreeding and long-term vegetative propagation [41].

One of the most widely used measures of genetic variation is nucleotide diversity (π), which quantifies the level of sequence polymorphism within a population or species (Table 2a,b). Nucleotide diversity is defined as the average number of nucleotide differences per site between two DNA sequences randomly sampled from a population. Unlike simple counts of polymorphic sites or single-nucleotide polymorphisms (SNPs), π incorporates both the frequency and distribution of sequence variants, thereby providing a measure of the average genetic divergence among individuals within a population. Comparative studies have shown that nucleotide diversity varies widely among species and populations, reflecting the combined influence of demographic history, reproductive biology, mutation rates, and selection [57].

$$\pi = \frac{\sum_{i<j} d_{ij}}{\binom{n}{2} L}$$

where:

$\sum_{i<j} d_{ij}$: The sum of the number of nucleotide differences between all possible pairs of sequences.
$\binom{n}{2}$: The total number of pairs.
L: the length of the DNA sequence.

Selfed PN40024 displayed the lowest levels of genetic variation among the groups analyzed, with a nucleotide diversity (π) of 0.0003 and an observed heterozygosity (Ho) of 0.01 [41]. These values reflect the extensive reduction in heterozygosity resulting from multiple generations of self-fertilization during the development of the reference genotype. In contrast, PN-derived clonal lineages and other cultivated European grapevine groups

retained substantially higher levels of diversity, with nucleotide diversity values around 0.0024 [41]. Interestingly, average heterozygosity was reported to be slightly higher in cultivated clonal lineages than in some wild populations, likely reflecting the long-term preservation of heterozygous genotypes through vegetative propagation.

Table 2. (**a**) Values of nucleotide diversity in different plant species. (**b**) Values of nucleotide diversity in different varieties of rice and grapes.

(**a**)

Species	**Reproduction Type**	**π ($\times 10^{-3}$)**	**Observation**
Arabidopsis thaliana [58]	Autogamous (selfing)	0.7–1.0	Very low due to self-pollination.
Oryza sativa [59]	Autogamous (mainly)	2.0–3.0	Severe bottleneck due to domestication.
Glycine max [60]	Autogamous	1.0–2.0	Limited due to domestication.
Vitis vinifera [45,61]	Alogamous (cultivated hermaf.)	7.29–13.9	Very high for a crop.
Zea mays [62]	Monoic (alogamous)	10.0–15.0	One of the highest levels among crops.
Populus spp. [63]	Dioecious (obligate alogamous)	10.0–16.0	Forest tree with high gene flow.
Silene (dioecious species) [64]	Dioecious	12.0–30.0	Used as a model of high diversity.

(**b**)

Group	**Species**	**Nucleotide Diversity (π) ($\times 1000$)**	**Source**
Wild Progenitor	*O. rufipogon*	9.2	Huang et al. (2012) [59]
Domesticated (Indica)	*O. sativa*	2.2	Huang et al. (2012) [59]
Pinot Noir clones	*V. vinifera*	2.4	Xiao et al. (2025) [41]
"Wild" ME grapes	*V. vinifera*	3.9	Xiao et al. (2025) [41]
Wild and cultivated grapes	*V. vinifera*	13.9–14.7	Zhou et al. (2017) [45]

The Middle Eastern (ME) wild group analyzed by Xiao et al. [41] exhibited a nucleotide diversity of 0.0039, higher than that observed in cultivated lineages but still relatively modest compared with values reported for some wild progenitor populations of other crop species. These patterns are consistent with the combined effects of domestication, demographic bottlenecks, recurrent selection, and extensive clonal propagation, all of which have contributed to shaping contemporary patterns of genetic diversity in grapevine [41].

Nevertheless, the levels of nucleotide diversity observed in *Vitis* are not unusual among perennial crops that have undergone extensive vegetative propagation. A comparable pattern is observed in *Citrus*, where many commercially important cultivars, including oranges, grapefruits, and lemons, originated through complex interspecific hybridization events involving the ancestral taxa mandarin (*Citrus reticulata*), pomelo (*Citrus maxima*), and citron (*Citrus medica*) [65]. Once favorable hybrid genotypes were established, their propagation relied predominantly on clonal reproduction, preserving highly heterozygous genomic constitutions across successive generations.

In both grapevine and citrus, vegetative propagation limits the genetic reshuffling associated with sexual reproduction and thereby maintains allelic combinations that would otherwise be broken apart by recombination. Over long periods, this process also permits the accumulation and fixation of somatic mutations within clonal lineages, contributing to intracultivar variation and the emergence of distinct clones. Consequently, clonal propagation plays a major role in preserving genetic diversity at the cultivar level and shaping the long-term evolutionary dynamics of perennial crop genomes.

Using contemporary demographic and population-genomic approaches, Xiao et al. [41] identified substantial signatures of shared ancestry and introgression between wild European (EU) populations and the modern Pinot Noir genome, particularly on chromosomes 1, 2, 3, and 19. Notably, some of these regions overlap genomic intervals associated with key domestication-related traits, including the sex-determination region on chromosome 2 and loci involved in fruit development and ripening on chromosome 19. The presence of introgressed haplotypes within these functionally important regions highlights the complex evolutionary interactions that have shaped grapevine genomes throughout domestication and post-domestication history.

However, the directionality of these introgression events remains difficult to establish unequivocally. Given the relatively modest nucleotide diversity observed in the sampled EU and Middle Eastern (ME) wild populations [41], one possible interpretation is that at least some shared genomic segments may reflect historical feralization processes, whereby cultivated alleles introgressed into wild or semi-wild populations through recurrent gene flow. Alternatively, these regions may represent the persistence of ancestral polymorphisms or introgression from wild populations into cultivated lineages. Distinguishing among these scenarios will require additional demographic modeling, broader population sampling, and the integration of archaeological and historical genomic data.

A notable methodological limitation of these comparisons is the lack of published assembly quality metrics, including BUSCO (Benchmarking Universal Single-Copy Orthologs) completeness scores, for the specific EU and ME wild populations analyzed. While the recently completed Pinot Noir T2T reference genome provides an exceptionally complete and highly resolved genomic framework, the absence of equivalent assembly and annotation quality assessments for the putative wild populations complicates direct comparisons among datasets. Differences in assembly completeness, haplotype resolution, or variant-calling pipelines may influence estimates of shared ancestry, introgression, and genetic diversity.

Consequently, caution is warranted when interpreting the magnitude and directionality of introgression signals. Without standardized quality metrics and harmonized analytical workflows, it remains challenging to determine the extent to which observed genomic patterns reflect genuine biological processes as opposed to technical artifacts arising from differences in genome assembly, sequence coverage, or bioinformatic filtering. Future studies incorporating fully characterized, high-quality assemblies from diverse wild *Vitis vinifera* subsp. *sylvestris* populations will be essential for resolving these uncertainties.

The strong genetic affinity observed between Pinot Noir and European wild grapevine populations—greater than that detected with geographically distant Middle Eastern groups—illustrates a recurring challenge in population genomics: the inference of gene-flow directionality. Current demographic models often interpret this pattern as evidence of localized post-domestication introgression from European wild populations into cultivated grapevines during their westward dispersal. However, the same genomic signal may also be compatible with alternative demographic scenarios, including varying degrees of cultivar-to-wild gene flow associated with feralization processes.

Population-structure analyses, including ADMIXTURE-based approaches, consistently identify distinct genetic clusters corresponding to cultivated grapevines and neighboring wild populations. Nevertheless, such clustering patterns alone do not unambiguously distinguish between ancient divergence, recurrent introgression, or more recent feralization events. Similar genetic structures can emerge under multiple demographic histories, particularly when long-term gene flow, population bottlenecks, and local adaptation have acted simultaneously. Consequently, the interpretation of shared ancestry between cultivated and

wild grapevine populations remains sensitive to the demographic assumptions underlying the analytical framework employed.

Discriminating among these alternative scenarios will likely require the integration of haplotype-resolved genomes, high-quality assemblies from geographically diverse *Vitis vinifera* subsp. *sylvestris* populations, and explicit demographic modeling capable of testing competing hypotheses of domestication, introgression, and feralization.

5. Beyond the Model: The Genetics of Domestication

5.1. Genetic Variation and Trait Identification

The transition from wild grapevine (*Vitis vinifera* subsp. *sylvestris*) to cultivated grapevine (*Vitis vinifera* subsp. *vinifera*) was accompanied by a suite of morphological, physiological, and genetic modifications collectively referred to as the domestication syndrome. The principal characteristics associated with this syndrome are summarized in Table 3 and include increased berry size, larger seeds, elevated sugar accumulation, and the transition from dioecious to predominantly hermaphroditic reproductive systems. In contrast, wild *sylvestris* populations typically produce smaller berries, smaller and more rounded seeds, higher-acidity fruits, and retain dioecious reproduction.

Table 3. Traits representative of the Domestication Syndrome.

Trait	Wild (*sylvestris*)	Domesticated (*vinifera*)
Sex	Dioecious (Male or Female)	Hermaphroditic (Self-pollinating)
Berry Size	Small (approx. 6 mm)	Large (up to 30 mm)
Berry Color	Mostly dark purple/black	Wide variety (white, red, black)
Sugar/Acidity	Lower sugar, higher acidity	Higher sugar, balanced acidity
Seed Shape	Round/Spherical	Pear-shaped (Pyriform)
Propagation	Seeds (Sexual)	Cuttings/Grafting (Clonal)

Among these traits, changes in seed morphology represent one of the most distinctive and extensively studied signatures of grapevine domestication. Archaeobotanical and morphometric analyses have consistently demonstrated that seed shape and size provide valuable markers for distinguishing wild and cultivated grapevine forms and for reconstructing the evolutionary history of domestication processes [66–69].

In recent years, we have investigated variation in grapevine seed morphology across a broad range of cultivars. Our approach is based on the representation of seed outlines using Elliptic Fourier Transform (EFT) analysis, which provides a quantitative framework for describing and comparing seed shape [70]. Fourier coefficients are extracted from digital images using the software MOMOCS 1.5.0. [71], after which average outlines and curvature profiles are generated using scripts implemented in *Mathematica* [72–74]. Curvature is quantified following methodologies originally developed for the analysis of *Arabidopsis* root morphology [75].

To date, morphological measurements, average outlines, and curvature parameters have been obtained for more than 200 cultivars, primarily originating from Central Europe and the Iberian Peninsula, together with dozens of *Vitis vinifera* subsp. *sylvestris* individuals from Spain, Portugal, and France. The dataset also includes six additional *Vitis* species, enabling a broader comparative analysis of seed morphology across the genus.

These analyses reveal four principal morphological groups: (I) *Vitis* species other than *V. vinifera* and *V. vinifera* subsp. *sylvestris*; (II) the 'Hebén'–'Furmint' morphotype; (III) the 'Traminer' morphotype; and (IV) the 'Muscat Hamburg'–'Königin der Weingärten' morphotype (Figure 3).

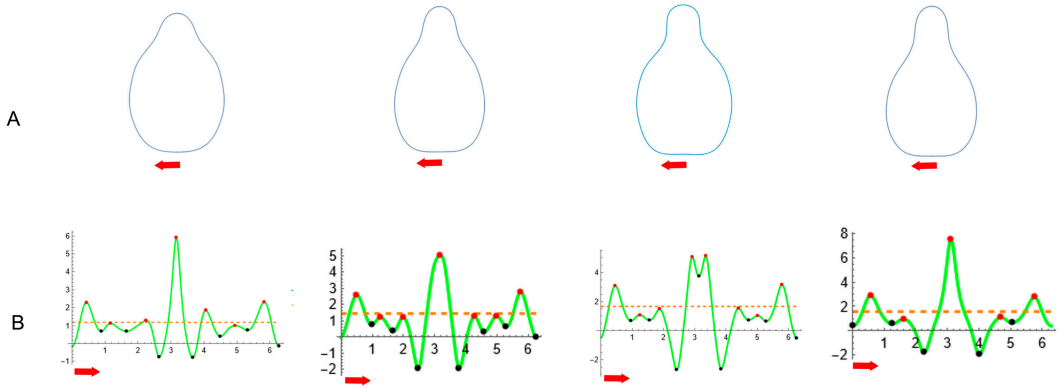

Figure 3. Morphological outlines and quantitative curvature analysis of *Vitis* seeds based on Elliptic Fourier Transforms (EFT). (**A**) Top panel: Closed seed contours reconstructed from the average Fourier coefficients utilizing 8 harmonics to filter high-frequency noise while preserving key diagnostic shapes. From left to right: *Vitis amurensis* Rupr., and three cultivars of *Vitis vinifera* L. (*vars* Hebén, Traminer, and Muscat Hamburg), arranged to illustrate the gradient of pyriform styling, stalk definition, and chalaza positioning. (**B**) Bottom panel: Corresponding curvature function plots modeled with Wolfram Mathematica, representing the variation in curvature (green) as a function of the perimeter distance mapped along the seed outline (0 to 2 π radians). The continuous green line represents the curvature trajectory. Notable geometric landmarks along the boundary are highlighted dynamically: red dots mark maximum convexities/peaks (positive curvature values), whereas black dots isolate inflections (two major inflections are the negative curvature values conserved in all four samples). The horizontal orange dashed line represents average curvature values. The red arrows indicate both the precise anatomical starting point (base of the seed) and the clockwise direction of the perimeter measurement.

The four morphological groups are distinguished by clear differences in seed geometry. Group I exhibits the highest solidity, whereas Group II displays intermediate values. In contrast, Groups III and IV are characterized by lower solidity and more pronounced outline curvature. Specifically, Group III shows the highest absolute negative curvature, while Group IV exhibits the highest positive curvature. Solidity, defined as the ratio between the projected seed area and the area of its convex hull, provides a quantitative measure of outline compactness. Mean solidity values for Groups I, II, III, and IV are approximately 0.968, 0.954, 0.946, and 0.945, respectively. Across the dataset, curvature and solidity are inversely related, with increasingly curved seed outlines corresponding to lower solidity values.

Interestingly, Groups II, III, and IV correspond closely to haplotypes 6, 5, and 3 identified by Dong et al. [76], represented by the cultivars Hebén, Traminer, and Muscat Hamburg, respectively. This correspondence suggests a potential association between seed morphology and major maternal lineages within cultivated grapevine.

Genome sequencing and comparative genomic analyses have substantially advanced our understanding of grapevine domestication, adaptation, and sexual differentiation. Zhou et al. [45] performed whole-genome sequencing of 18 individuals representing 14 *Vitis vinifera* cultivars, nine wild *V. vinifera* subsp. *sylvestris* accessions, and one *V. rotundifolia* outgroup, identifying genomic signatures associated with domestication and differentiation between wine and table grapes. Candidate genes under selection included loci involved in sugar transport and metabolism, notably the SWEET1 sugar transporter, as well as genes associated with flower development, berry development, and stress responses.

Compared with wild populations, cultivated grapevines were found to harbor approximately 5.2% more putatively deleterious variants, with a greater proportion maintained in the heterozygous state. This pattern is consistent with the long-term effects of clonal propagation, which can preserve mildly deleterious alleles by limiting their exposure to

purifying selection through sexual recombination [45]. In contrast, wild grapevines exhibited an enrichment of genes involved in environmental adaptation and defense, including pathways related to flavonoid biosynthesis, ethylene signaling, and stilbenoid production, all of which contribute to responses against biotic and abiotic stressors.

Comparative genomic analyses have also highlighted important modifications within the sex-determination region on chromosome 2, a key locus underlying the transition from dioecy in wild populations to predominantly hermaphroditic reproduction in cultivated grapevines [41]. More broadly, genome-wide comparisons among cultivated varieties have enabled the identification of genetic variants associated with ripening time, cluster compactness, berry morphology, fruit composition, and numerous other agronomically important traits [77–79].

Furthermore, haplotype-resolved assemblies of cultivars such as Chardonnay and Nebbiolo have revealed extensive structural variation between homologous chromosomes, including large insertions, deletions, inversions, and other rearrangements that were largely undetectable using earlier short-read sequencing approaches [80,81].

Most North American *Vitis* species exhibit natural resistance to grape phylloxera (*Daktulosphaira vitifoliae*), making them invaluable resources for rootstock breeding. Among these, *Vitis riparia* and *Vitis rupestris* combine phylloxera resistance with excellent rooting ability from hardwood cuttings and good graft compatibility with *V. vinifera* scions. *Vitis berlandieri* contributes additional resistance to phylloxera together with tolerance to calcareous soils. Consequently, many of the rootstocks adopted worldwide were developed through crosses between *V. berlandieri* and either *V. riparia* or *V. rupestris*, combining lime tolerance with improved rooting performance.

Recent advances in high-fidelity (HiFi) long-read sequencing have enabled the generation of chromosome-scale, diploid genome assemblies for major rootstocks, including '110R', 'Kober 5BB', and '101-14 Mgt' [82,83]. These assemblies have facilitated the precise identification of parental genomic contributions and have improved our understanding of the genetic basis of resistance to phylloxera and other environmental stresses.

Comparative genomics has also revealed extensive structural changes associated with grapevine evolution and domestication. Long et al. [84] analyzed 17 chromosome-scale grapevine assemblies, using *Vitis rotundifolia* as an outgroup and comparing cultivated grapevines, wild progenitors, and wild relatives. Their results indicated that variation in genome size among lineages is largely attributable to differences in transposable element (TE) content. Long terminal repeat (LTR) retrotransposons of the Gypsy and Copia superfamilies have played a major role in shaping genome architecture and promoting lineage-specific genome expansion.

In North American species such as *Vitis labrusca*, these transposable elements appear to have contributed to the expansion and diversification of disease-resistance loci, including nucleotide-binding leucine-rich repeat (NLR) gene clusters located on chromosomes 12 and 18 [85]. Such findings highlight the importance of repetitive DNA not only as a driver of genome evolution but also as a source of genetic variation underlying adaptive traits and disease resistance.

The growing number of phased telomere-to-telomere (T2T) genome assemblies, together with increasingly sophisticated pipelines for high-confidence single-nucleotide polymorphism (SNP) detection and structural variant (SV) discovery, has greatly expanded our understanding of adaptive evolution in the genus *Vitis*. These resources have been instrumental in identifying the genetic basis of important survival traits, including cold tolerance in *Vitis amurensis* Rupr. [86] and disease resistance mechanisms in *Vitis davidii* (Rom.Caill.) Foëx and *Vitis riparia* A. Gray [87,88].

One major consequence of this rapid expansion in genome resources has been the development of grapevine pangenomes. By integrating genomic information from cultivated varieties, wild relatives, and interspecific hybrids, pangenomic frameworks capture genetic variation that is absent from any single reference genome. These resources have proven valuable for identifying genes and structural variants associated with agronomically important traits, thereby supporting breeding programs aimed at improving productivity, resilience, and environmental sustainability. North American *Vitis* species have long served as reservoirs of adaptive alleles, contributing traits such as salt tolerance and resistance to *Xylella fastidiosa*, downy mildew, and other pathogens [89].

Pangenomic and haplotype-resolved analyses have also facilitated the characterization of unique cultivar phenotypes. For example, the cultivar 'Julian' (JA), which simultaneously bears flowers and berries at multiple developmental stages, has been used to identify candidate genes and structural variants potentially involved in the regulation of floral transition and developmental timing [40].

At a finer scale, advances in variant-detection methodologies are transforming the study of clonal diversity. Because genetic differences among clones are often limited to a small number of somatic mutations, approaches such as DeepVariant and k-mer–based analyses are increasingly employed to detect rare single-nucleotide variants (SNVs) and small insertions/deletions (indels) that might otherwise be filtered out as sequencing errors. By exploiting the high accuracy of PacBio HiFi sequencing, these methods enable the reliable identification of clonal polymorphisms and provide new insights into the morphological and physiological divergence that accumulates during long-term vegetative propagation [90].

Furthermore, comparative analyses of DNA obtained from different tissues, such as roots and leaves, have improved the detection of chimeric genotypes, an important source of intra-cultivar variation. The identification and characterization of chimeras provides valuable opportunities for clonal selection and the development of improved grapevine cultivars [91].

5.2. Origins of Domestication and Diversification

Magris et al. [61] conducted a comprehensive analysis of 204 *Vitis vinifera* genomes, identifying 7,364,288 SNPs within non-repetitive genomic regions. This high-resolution dataset enabled the assignment of individual accessions to distinct geographic and genetic groups, broadly corresponding to the eco-geographical classifications originally proposed by Negrul, including the *orientalis*, *pontica georgica*, *pontica balcanica*, and *occidentalis* groups. These results provide strong genomic support for the historical framework established through classical ampelographic and geographical observations.

At the same time, the study highlights ongoing challenges in reconstructing grapevine domestication history, particularly with regard to distinguishing truly wild *Vitis vinifera* subsp. *sylvestris* populations from feral lineages derived from cultivated grapevines. Because recurrent gene flow, introgression, and feralization can generate similar genomic signatures, the interpretation of population structure and ancestry remains dependent on both sampling strategies and the demographic models employed.

Magris et al. [61] reported a high degree of shared genetic variation between wild and cultivated grapevines, with only 8.2% of SNPs identified in *V. vinifera* subsp. *sylvestris* being absent from cultivated accessions. Such extensive overlap is consistent with a history of recurrent gene flow between wild and cultivated populations and/or a relatively moderate domestication bottleneck. Similarly, Myles et al. [92] interpreted the substantial levels of diversity retained in cultivated grapevines as evidence that domestication was not accompanied by a severe reduction in genetic variation.

Nevertheless, the magnitude of the domestication bottleneck in grapevine remains debated, largely because the distinction between genuinely wild and feral populations is often difficult to establish. As emphasized by Terral et al. [69], naturalized vines and continuous gene flow from vineyards into riparian ecosystems can obscure the genetic boundaries between wild and cultivated compartments. Consequently, estimates of ancestral diversity may be sensitive to the composition of the reference populations used in genomic analyses.

This issue is particularly relevant when interpreting comparisons between cultivated grapevines and putatively wild *sylvestris* populations. If some accessions classified as wild possess partial feral ancestry, levels of genetic differentiation between wild and cultivated groups may be underestimated, potentially influencing demographic inferences regarding bottleneck intensity and historical gene flow. In this context, the relatively low nucleotide diversity reported for some wild populations ($\pi \approx 3.8 \times 10^{-3}$) compared with cultivated grapevines [61] may reflect a combination of demographic history, sampling effects, introgression, and varying degrees of feralization.

Furthermore, the long-term use of vegetative propagation in cultivated grapevines has contributed to the preservation of heterozygous genomic constitutions and the accumulation of somatic mutations within clonal lineages. As a result, cultivated populations can retain substantial genetic diversity despite the demographic constraints typically associated with domestication.

This interpretation is further informed by the f3-statistics reported by Magris et al. [61]. The strongest signals of shared ancestry involve primitive European cultivars and French populations classified as *V. vinifera* subsp. *sylvestris*, with Z-scores reaching highly significant negative values (up to −43.7). Such results indicate extensive shared genetic history and are consistent with substantial admixture or recurrent gene flow between cultivated and wild compartments.

Notably, the strongest signals are observed in Western European populations, whereas weaker values are reported for populations from the Caucasus (e.g., Z = −9.8), a region frequently considered part of the primary domestication area. Although these patterns have commonly been interpreted as evidence of localized introgression between cultivated grapevines and neighboring wild populations, alternative demographic scenarios remain plausible. One such possibility is that some populations currently classified as wild may contain varying degrees of feral ancestry derived from historical cultivated lineages.

This interpretation is supported by the observation that similar statistical patterns are obtained when Western *sylvestris* populations are replaced by Western feral populations in Supplementary Table 1 of Magris et al. [61]. While this result does not demonstrate a feral origin for the French populations, it highlights the difficulty of distinguishing among ancient wild ancestry, recurrent introgression, and feralization using contemporary genomic data alone. Additional evidence from haplotype-resolved genomes, demographic modeling, and carefully validated wild reference populations will be required to discriminate among these competing scenarios.

The widespread occurrence of shared haplotypes, including H4 and H7, across wild, feral, and cultivated grapevine populations, as reported by Magris et al. [61], underscores the extensive genetic continuity linking these compartments. Such patterns suggest that the genomic structure of Western European grapevines cannot be explained solely by simple models of localized introgression into an originally Eastern domesticated lineage. Instead, the observed haplotype sharing points to a more complex evolutionary history involving repeated episodes of gene flow, population mixing, and demographic restructuring.

At least two non-mutually exclusive scenarios may account for these observations. The first invokes the contribution of an ancient, genetically differentiated European *sylvestris* lineage, potentially representing an unsampled or extinct population that participated in

secondary domestication processes. The second emphasizes the role of long-term feralization, whereby cultivated grapevines repeatedly escaped cultivation, established naturalized populations, and subsequently exchanged genes with neighboring wild populations. Under this scenario, part of the genetic signal currently attributed to ancient wild ancestry could instead reflect the persistence of diversity originating from historical cultivated gene pools that are no longer represented among modern commercial cultivars.

Regardless of the relative importance of these mechanisms, the extensive haplotype sharing observed among wild, feral, and cultivated grapevines highlights the limitations of strictly binary classifications of grapevine populations. The long history of vegetative propagation has preserved highly heterozygous genomic constitutions within cultivated lineages, while recurrent introgression and feralization have generated complex patterns of ancestry across the species range. Consequently, contemporary grapevine populations are best viewed as components of a dynamic evolutionary continuum rather than as discrete "wild" and "domesticated" entities [61].

An important limitation of the dataset analyzed by Magris et al. [61] is the limited representation of Iberian germplasm, particularly from Spain and Portugal. This geographic gap is noteworthy because the Iberian Peninsula has long been recognized as a major center of grapevine diversification and a key contact zone linking Mediterranean, North African, and Western European genetic lineages. Consequently, incomplete sampling of Iberian cultivars may reduce the resolution with which demographic relationships among Western European grapevine populations can be reconstructed.

The absence of cultivars descended from Hebén is especially relevant in this context. Recent pedigree analyses [32,33] have identified Hebén as a major founder of the Iberian grapevine gene pool, contributing ancestry to more than 60 cultivars, including Cayetana Blanca, Pedro Ximénez, and Airén. Given its central position within Iberian pedigree networks, the inclusion of Hebén and its descendants could provide valuable information regarding the historical relationships among cultivated grapevines, local *sylvestris* populations, and putative feral lineages.

From a population-genomic perspective, the absence of this substantial component of Western European diversity may influence estimates of population structure and ancestry. A broader sampling of Iberian germplasm could reveal additional intermediate haplotypes and shared ancestry patterns linking cultivated and wild populations across Western Europe. Whether such data would modify the distinction between cultivated and wild groups identified by Magris et al. [61] remains an open question, but their inclusion would undoubtedly provide a more comprehensive framework for evaluating alternative models of domestication, introgression, and feralization.

These findings can be further interpreted in the context of the dual-domestication model proposed by Dong et al. [76], which suggests that grapevine domestication involved at least two major domestication trajectories centered in the Caucasus and the Levant. According to this model, Western wine grapes derive much of their ancestry from the Caucasian lineage, but their subsequent dispersal across Europe was accompanied by recurrent hybridization and introgression with local *Vitis vinifera* subsp. *sylvestris* populations. As a result, modern European grapevine genomes retain signatures of both long-range migration and regional admixture.

Additional support for the complexity of grapevine evolution comes from comparative pangenomic analyses conducted by Long et al. [84]. By comparing multiple chromosome-scale assemblies representing cultivated grapevines, wild progenitors, and related species, the authors documented extensive patterns of gene gain, gene loss, and structural variation. These results indicate that domestication was accompanied by substantial genomic reorga-

nization affecting pathways involved in environmental adaptation, disease resistance, and fruit development.

Collectively, these findings highlight the limitations of interpreting grapevine evolutionary history through rigid ancestry categories alone. Recurrent introgression, feralization, and long-term vegetative propagation have generated genomes that frequently contain genetic components associated with both wild and cultivated populations. Consequently, the distinction between "wild" and "domesticated" grapevines is often less discrete than implied by simple classification schemes, emphasizing the need for demographic models capable of accommodating continuous gene flow and complex patterns of shared ancestry.

5.3. Feralization, Wild Introgression, and the "Feral Problem"

In practice, *Vitis vinifera* subsp. *sylvestris* is often identified by a combination of ecological, morphological, and reproductive characteristics rather than by a single diagnostic genetic marker. Wild status is typically inferred from three principal criteria: (1) occurrence in natural habitats largely isolated from commercial viticulture, (2) characteristic seed morphology, generally involving smaller and more rounded pips, and (3) dioecious reproduction. Consequently, *sylvestris* is frequently defined through a suite of correlated traits rather than through the presence of a unique "wild" genotype.

The reproductive criterion is particularly noteworthy considering recent advances in the genetics of sex determination. The transition from dioecy in wild grapevines to predominantly hermaphroditic flowers in cultivated forms is controlled by a relatively small genomic region on chromosome 2. This finding complicates the interpretation of reproductive phenotype as an absolute indicator of ancestry, since recurrent gene flow between wild and cultivated populations can potentially decouple reproductive traits from broader genomic backgrounds. As a result, individuals occupying intermediate ecological settings may exhibit combinations of morphological, reproductive, and genetic characteristics that are difficult to classify unequivocally as either wild or cultivated.

These observations highlight the challenges of distinguishing remnant wild populations from feral or introgressed lineages using any single criterion alone. A robust identification of authentic *sylvestris* populations increasingly requires the integration of ecological information, reproductive biology, morphometric analyses, and genome-wide data capable of resolving both nuclear and cytoplasmic ancestry.

In grapevine, sexual phenotype is primarily determined by a small, highly differentiated region on chromosome 2 known as the Sex Determination Region (SDR) [37]. This region functions as a supergene, comprising tightly linked genetic factors that control male and female reproductive functions. Recombination within the SDR is strongly suppressed, allowing specific combinations of sex-determining alleles to be inherited as stable haplotypes across generations [37].

Three principal haplotypes have been identified within the SDR, corresponding to the male (M), hermaphrodite (H), and female (F) forms. These haplotypes exhibit a hierarchical dominance relationship:

- M (male haplotype)—dominant
- H (hermaphrodite haplotype)—intermediate dominance
- F (female haplotype)—recessive

Therefore, male plants typically possess at least one M haplotype (MF or MH), hermaphroditic plants carry the H haplotype in the absence of M (HF or HH), and female plants are generally homozygous for the recessive F haplotype (FF) [37]. This simple genetic architecture explains the inheritance of sexual phenotype in both wild and cultivated grapevine populations and provides a mechanistic framework for understanding the

transition from dioecy in *Vitis vinifera* subsp. *sylvestris* to predominantly hermaphroditic reproduction in cultivated grapevines.

According to this dominance hierarchy, wild female grapevines are typically homozygous for the recessive female haplotype (FF), whereas wild males generally carry at least one dominant male haplotype (most commonly MF) [37]. In contrast, cultivated hermaphroditic grapevines are usually HF or HH. This genetic architecture has important implications for the interpretation of feral populations.

A population derived exclusively from escaped cultivated grapevines would initially contain only H and F haplotypes. Under such circumstances, offspring produced through self-fertilization or intercrossing among feral descendants would be expected to segregate into hermaphroditic (HH, HF) and female (FF) phenotypes but would not generate male genotypes because the dominant M haplotype is absent from the founding population [37]. Consequently, the long-term establishment of fully dioecious populations from cultivated lineages alone would be genetically constrained.

The situation changes once gene flow from wild populations is introduced. Pollen originating from male *Vitis vinifera* subsp. *sylvestris* individuals can reintroduce the M haplotype into feral or introgressed populations, generating male offspring and restoring the genetic basis for dioecy. Therefore, the presence of male plants within mixed or disturbed habitats may provide evidence of historical or ongoing genetic exchange with wild grapevine populations, although the extent and timing of such introgression require verification through genome-wide analyses.

The transition from wild *Vitis vinifera* subsp. *sylvestris* to cultivated *V. vinifera* subsp. *vinifera* represents far more than a phenotypic shift in berry size, seed morphology, or floral architecture; it reflects a profound genomic transformation associated with domestication, selection, and long-term human management. Myles et al. [92] provided one of the first comprehensive genome-wide analyses of this transition, revealing both the substantial diversity retained in cultivated grapevines and the extensive genetic continuity linking wild and cultivated populations.

Despite these advances, the identification of genuinely wild, minimally introgressed *sylvestris* populations remains a major challenge. Ecological and genetic studies in the Rhine Valley have questioned whether indigenous *sylvestris* lineages persist as distinct populations or survive only as genetic components within naturalized and introgressed grapevine communities [93]. Similar conclusions have emerged from surveys conducted in Italy [39] and the Iberian Peninsula [38], where recurrent gene flow between cultivated and wild compartments has generated complex patterns of admixture that obscure population boundaries.

These observations underscore the importance of integrating ecological, morphological, and genomic criteria when characterizing wild grapevine populations. Studies in regions such as the Donau-Auen [93], Tuscany [39], and the Iberian Peninsula [38] have emphasized the value of sampling geographically isolated populations and incorporating reproductive traits, particularly dioecy, into identification protocols. Additional evidence from seed morphometrics, chloroplast haplotypes, and genome-wide markers can further improve the discrimination of remnant wild populations from feral or introgressed lineages.

Consequently, no single criterion is sufficient to establish wild status with confidence. Rather, the identification of authentic *V. vinifera* subsp. *sylvestris* populations increasingly relies on a multidisciplinary framework that combines reproductive biology, morphology, ecology, and high-resolution genomic analyses.

Population-genetic analyses indicate that feral grapevines do not simply revert to an ancestral wild state following escape from cultivation. Instead, they frequently participate in complex processes of de-domestication, maintaining genetic signatures inherited from

cultivated ancestors while simultaneously acquiring alleles through introgression with local wild populations. High-density SNP genotyping and whole-genome sequencing have revealed that many naturalized grapevine populations retain substantial levels of genome-wide heterozygosity, reflecting both their cultivated origins and their subsequent demographic histories [94].

Through recurrent hybridization with neighboring *Vitis vinifera* subsp. *sylvestris* populations, feral lineages can act as genetic bridges linking wild and cultivated gene pools. These interactions generate mosaics of ancestry that complicate the interpretation of population structure and challenge simple binary classifications of grapevine populations. Rather than representing entirely discrete entities, wild, feral, and cultivated grapevines often form interconnected evolutionary networks shaped by continuous gene flow, local adaptation, and historical demographic processes.

Recent advances in comparative genomics further emphasize the dynamic nature of grapevine genome evolution. Using 29 genome assemblies representing multiple *Vitis* species, Wang et al. [95] demonstrated that Helitron transposable elements have played a major role in the expansion and regulation of nucleotide-binding leucine-rich repeat (NLR) immune-receptor genes. Their analyses identified structural variants associated with resistance to downy mildew and provided new phylogenomic insights into the evolution of disease-resistance mechanisms across the genus. These findings illustrate how transposable elements contribute directly to adaptive variation by reshaping genome structure, modifying gene expression, and influencing the distribution of resistance traits among wild and cultivated grapevines.

More broadly, such studies reinforce the importance of structural variation and repetitive DNA in grapevine evolution. Once considered largely genomic "dark matter," transposable elements are now recognized as major drivers of genome diversification, adaptation, and ecological resilience throughout the genus *Vitis* [95].

6. Applications in Taxonomy

The transition from classical ampelography to high-resolution genomics has transformed grapevine taxonomy by providing unprecedented power to resolve both ancestral relationships and intra-cultivar variation. Two major applications have emerged from this genomic revolution: parentage reconstruction and clonal differentiation.

6.1. Parentage Verification and Pedigree Reconstruction

While traditional ampelography inferred relationships primarily from morphological similarities, genome-wide SNP analyses provide direct genetic evidence of ancestry. One of the most celebrated examples is the confirmation that Cabernet Sauvignon originated from a spontaneous cross between Cabernet Franc and Sauvignon Blanc [28]. Since then, high-density molecular markers and whole-genome sequencing have enabled the reconstruction of increasingly complex multi-generational pedigrees across the grapevine germplasm.

These approaches have revealed the pivotal role of founder cultivars in shaping regional diversity. In the Iberian Peninsula, for example, pedigree analyses have identified Hebén as a major ancestral cultivar contributing to a large number of modern varieties [32,33]. The genetic evidence supporting this central position is consistent with morphometric analyses that highlighted the distinctive characteristics of the Hebén lineage [31]. Beyond clarifying cultivar origins, pedigree reconstruction facilitates the identification of ancestral genotypes that lie outside currently recognized family networks and may harbor unique allelic combinations of value for conservation and breeding programs.

6.2. Clonal Identification: The Final Frontier

A longstanding limitation of traditional grapevine taxonomy has been its inability to distinguish among clones, which originate through the accumulation of somatic mutations within a single cultivar lineage. Although clones may exhibit important differences in agronomic performance, berry composition, disease susceptibility, or phenology, they are often indistinguishable using conventional morphological descriptors.

Recent advances in long-read sequencing (LRS) and haplotype-resolved genome assembly have substantially improved the detection of clonal variation. These technologies allow the identification of the specific single-nucleotide variants (SNVs), insertions/deletions (indels), and structural variants (SVs) that distinguish individual clones. Comparative genomic analyses have shown that vegetative propagation is accompanied by the gradual accumulation of somatic mutations, generating lineage-specific genomic signatures that can now be characterized with high confidence [96].

As sequencing technologies continue to improve, clonal identification is increasingly moving from a descriptive discipline to a precise genomic science. The ability to characterize clone-specific variation not only refines taxonomic classification but also provides new opportunities for germplasm management, clonal selection, and the preservation of valuable genetic diversity within traditional cultivars.

6.3. Bridging the Phenotype-Genotype Gap

The integration of genomic, biochemical, and morphometric approaches has substantially strengthened grapevine taxonomy by linking observable phenotypes to their underlying genetic determinants. This convergence has enabled the transition from purely descriptive classifications toward a more mechanistic understanding of varietal diversity.

Chemotaxonomy: Secondary metabolites provide valuable taxonomic markers that complement both morphological and genomic analyses. A notable example is anthocyanin composition, which reliably distinguishes *Vitis vinifera* from many North American *Vitis* species. Cultivated *V. vinifera* typically accumulates anthocyanin mono-glucosides, whereas American species frequently contain substantial proportions of anthocyanin diglucosides [97]. This chemotaxonomic distinction, first recognized by Ribéreau-Gayon [98], has subsequently been linked to specific genetic mechanisms controlling berry pigmentation. In particular, structural variation and transposable-element insertions affecting the regulatory gene *VvMybA1* play a central role in determining anthocyanin accumulation and berry color variation among cultivars [78].

Numerical Taxonomy: Advances in quantitative morphometrics have transformed classical ampelography from a largely descriptive discipline into a statistically robust analytical framework [22,23]. Modern image-analysis methods allow the objective quantification of shape variation in leaves, seeds, berries, and other organs, facilitating comparisons among cultivars and species. For example, Chitwood et al. [99] developed a high-resolution leaf morphospace integrating both genetic and developmental components of leaf architecture. By combining geometric morphometrics with predictive modeling, this approach enables the classification of grapevine leaves while simultaneously disentangling the respective contributions of genotype, developmental stage, and environmental influences to leaf morphology.

Together, chemotaxonomic, morphometric, and genomic approaches provide complementary perspectives on grapevine diversity. Their integration offers a powerful framework for resolving taxonomic relationships, identifying cultivars, and linking phenotypic variation to its molecular basis.

Palynology: Micro-morphological analysis of pollen grains remains a vital diagnostic tool for differentiating wild lineages from domesticated grapevine populations,

highlighting a conserved pollen structure in *V. sylvestris*. Micro-morphological analysis leverages sexual dimorphism—specifically tracking the shift from strictly dioecious, inaperturate/prolate wild structures to uniformly hermaphroditic, 3-colporate cultivated forms—to distinguish true wild germplasm from naturalized feral escapees [100].

Seed Morphometrics & Genomics: Seed morphology is profoundly altered by domestication, exhibiting a phenotypic shift distinct enough to differentiate wild sylvestris progenitors from domesticated varieties, as well as to discriminate among individual cultivars [74,101]. Digital seed analysis utilizing Elliptic Fourier Descriptors (EFDs) has become central to identifying specific cultivars within the archaeological record [69]. To connect these physical traits to underlying genomic regions, modern high-throughput mapping technologies are now deployed. For instance, Single Primer Enrichment Technology (SPET) was recently utilized to dissect the genetic architecture of discrete traits like berry texture and defective seed development [102]. Because morphometric parameters such as aspect ratio, solidity, and curvature outlines are complex **polygenic traits** controlled by multiple loci, SPET's high-throughput capability to sequence thousands of targeted, genome-wide markers makes it an ideal platform to construct high-density maps and resolve the precise quantitative trait loci (QTLs) governing these shape architectures.

7. Conclusions: The Future of *Vitis* Systematics

The evolution of *Vitis* taxonomy over the last two centuries reflects a profound transformation in botanical science. The discipline has progressed from the descriptive foundations established by pioneers such as Simon de Rojas Clemente y Rubio and Pierre Galet, who relied primarily on morphological observations, to an era in which whole-genome sequencing, pangenomics, and high-resolution phenotyping provide unprecedented insight into grapevine diversity, ancestry, and adaptation.

Advances in molecular genetics have fundamentally reshaped our understanding of grapevine evolution. Genome-wide analyses have revealed the complex demographic processes underlying domestication, diversification, introgression, and feralization, while haplotype-resolved and telomere-to-telomere assemblies have exposed the extensive structural variation hidden within the grapevine genome. These developments have demonstrated that no single reference genome can fully represent the diversity of the genus. Instead, the emerging pangenomic framework recognizes *Vitis* as a dynamic collection of core and accessory genomic components that vary among species, populations, cultivars, and even clones.

At the same time, the integration of genomics with ampelography, seed morphometrics, palynology, chemotaxonomy, and quantitative phenotyping has transformed taxonomy into a multidisciplinary science capable of linking genotype, phenotype, and evolutionary history. This synthesis has not only refined cultivar identification and pedigree reconstruction but has also provided new tools for understanding the origins and maintenance of genetic diversity across wild, feral, and cultivated grapevine populations.

The continued expansion of high-quality genome assemblies, pangenomes, and comparative phylogenomic resources will further improve our ability to reconstruct the evolutionary history of grapevines and identify adaptive genetic variation. Such knowledge is increasingly important in the context of climate change, emerging pathogens, and the growing demand for sustainable viticulture. Wild relatives, feral populations, and underutilized traditional cultivars represent valuable reservoirs of genetic diversity that may contribute traits associated with stress tolerance, disease resistance, and environmental resilience.

Given the breadth and rapid development of the field, this review cannot address every aspect of contemporary grapevine systematics in detail. Nevertheless, it highlights the major conceptual and technological advances that have transformed the discipline

and provides a framework for future investigations into the evolution, taxonomy, and conservation of the genus *Vitis*. The future of grapevine systematics will depend not on replacing classical approaches, but on integrating them with increasingly sophisticated genomic tools to achieve a more complete understanding of one of humanity's most important cultivated plants.

Author Contributions: Conceptualization, E.C. and J.L.R.L.; methodology, E.C. and J.L.R.L.; software, E.C. and J.L.R.L.; validation, E.C. and J.L.R.L.; formal analysis, E.C. and J.L.R.L.; investigation, E.C. and J.L.R.L.; resources, E.C. and J.L.R.L.; data curation, E.C. and J.L.R.L.; writing—original draft preparation, E.C. and J.L.R.L.; writing—review and editing, E.C. and J.L.R.L.; visualization, E.C. and J.L.R.L.; supervision, E.C. and J.L.R.L.; project administration, E.C. and J.L.R.L.; funding acquisition, E.C. and J.L.R.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations and Technical Terms

Allopolyploidy	Hybridization followed by chromosome doubling (relevant for wild hybrids).
BUSCO score	Benchmarking Universal Single-Copy Orthologs
DeepVariant	A bioinformatic tool using deep learning to call genetic variants with high accuracy, essential for detecting subtle clonal mutations.
Haplotype Phasing	The process of separating maternal and paternal genetic information.
k-mer Analysis	A method for comparing genomic sequences by breaking them into shorter strings of length k, useful for species differentiation without a reference genome.
NLR Genes	A large gene family involved in plant immunity; their distribution is a key taxonomic marker for wild *Vitis* species.
Presence-Absence Variation (PAV)	Structural variations where entire genes are present in one cultivar but absent in another.
Inbreeding Depression	The creation of PN40024 as a technical feat due to the general loss of vigor in selfed grapevines.

References

1. Herrera, F.; Carvalho, M.R.; Stull, G.W.; Jaramillo, C.; Manchester, S.R. Cenozoic Seeds of Vitaceae Reveal a Deep History of Extinction and Dispersal in the Neotropics. *Nat. Plants* **2024**, *10*, 1091–1099. [CrossRef] [PubMed]
2. Wen, J.; Lu, L.; Nie, Z.; Liu, X.; Zhang, N.; Ickert-Bond, S.; Gerrath, J.; Manchester, S.R.; Boggan, J.; Chen, Z. A New Phylogenetic Tribal Classification of the Grape Family (Vitaceae). *J. Syst. Evol.* **2018**, *56*, 262–272. [CrossRef]
3. Zeng, L.; Zhang, N.; Zhang, Q.; Endress, P.K.; Huang, J.; Ma, H. Resolution of Deep Eudicot Phylogeny and Their Temporal Diversification Using Nuclear Genes from Transcriptomic and Genomic Datasets. *New Phytol.* **2017**, *214*, 1338–1354. [CrossRef] [PubMed]
4. Zhang, N.; Wen, J.; Zimmer, E.A. Congruent Deep Relationships in the Grape Family (Vitaceae) Based on Sequences of Chloroplast Genomes and Mitochondrial Genes via Genome Skimming. *PLoS ONE* **2015**, *10*, e0144701. [CrossRef] [PubMed]
5. Wan, Y.; Schwaninger, H.R.; Baldo, A.M.; Labate, J.A.; Zhong, G.-Y.; Simon, C.J. A Phylogenetic Analysis of the Grape Genus (*Vitis* L.) Reveals Broad Reticulation and Concurrent Diversification during Neogene and Quaternary Climate Change. *BMC Evol. Biol.* **2013**, *13*, 141. [CrossRef] [PubMed]
6. Manchester, S.R.; Kapgate, D.K.; Wen, J. Oldest Fruits of the Grape Family (Vitaceae) from the Late Cretaceous Deccan Cherts of India. *Am. J. Bot.* **2013**, *100*, 1849–1859. [CrossRef] [PubMed]
7. Plants of the World Online. *Vitis* L. Kew Science. Available online: http://powo.science.kew.org/taxon/urn:lsid:ipni.org:names:325876-2 (accessed on 24 April 2026).
8. Renner, S.S.; Ricklefs, R.E. Dioecy and its correlates in the flowering plants. *Am. J. Bot.* **1995**, *82*, 596–606. [CrossRef]

9. Flora of North America Editorial Committee (Ed.) *Flora of North America: North of Mexico*; Oxford University Press: New York, NY, USA, 1993.
10. Mullins, M.G.; Bouquet, A.; Williams, L.E. *Biology of the Grapevine*; Cambridge University Press: Cambridge, UK, 1992.
11. Péros, J.-P.; Launay, A.; Peyrière, A.; Berger, G.; Roux, C.; Lacombe, T.; Boursiquot, J.-M. Species Relationships within the Genus *Vitis* Based on Molecular and Morphological Data. *PLoS ONE* **2023**, *18*, e0283324. [CrossRef] [PubMed]
12. Tröndle, D.; Schröder, S.; Kassemeyer, H.-H.; Kiefer, C.; Koch, M.A.; Nick, P. Molecular Phylogeny of the Genus *Vitis* (Vitaceae) Based on Plastid Markers. *Am. J. Bot.* **2010**, *97*, 1168–1178. [CrossRef] [PubMed]
13. Viala, P. *Une Mission Viticole en Amerique*; C. Coulet: Montpellier, France, 1889.
14. Ekhvaia, J.; Gurushidze, M.; Blattner, F.R.; Akhalkatsi, M. Genetic Diversity of *Vitis vinifera* in Georgia: Relationships between Local Cultivars and Wild Grapevine, *V. vinifera* L. subsp. *sylvestris*. *Genet. Resour. Crop Evol.* **2014**, *61*, 1507–1521. [CrossRef]
15. Riaz, S.; De Lorenzis, G.; Velasco, D.; Koehmstedt, A.; Maghradze, D.; Bobokashvili, Z.; Musayev, M.; Zdunic, G.; Laucou, V.; Walker, M.A.; et al. Genetic Diversity Analysis of Cultivated and Wild Grapevine (*Vitis vinifera* L.) Accessions around the Mediterranean Basin and Central Asia. *BMC Plant Biol.* **2018**, *18*, 137. [CrossRef] [PubMed]
16. Marrano, A.; Grzeskowiak, L.; Moreno Sanz, P.; Lorenzi, S.; Prazzoli, M.L.; Arzumanov, A.; Amanova, A.; Failla, O.; Maghradze, D.; Grando, M.S. Genetic Diversity and Relationships in the Grapevine Germplasm Collection from Central Asia. *Vitis* **2015**, *54*, 233–237.
17. Ricciardi, V.; Gherardi, M.; Cola, G.; Failla, O.; Maghradze, D.; Lorenzis, G.D. Georgian Wild Grapevine Germplasm Is Confirmed to Be Genetically Distinct from Mediterranean Germplasm. *OENO One* **2026**, *60*. [CrossRef]
18. De Lorenzis, G.; Mercati, F.; Bergamini, C.; Cardone, M.F.; Lupini, A.; Mauceri, A.; Caputo, A.R.; Abbate, L.; Barbagallo, M.G.; Antonacci, D.; et al. SNP Genotyping Elucidates the Genetic Diversity of Magna Graecia Grapevine Germplasm and Its Historical Origin and Dissemination. *BMC Plant Biol.* **2019**, *19*, 7. [CrossRef] [PubMed]
19. Rojas Clemente y Rubio, S. *Ensayo Sobre las Variedades de la Vid Común que Vegetan en Andalucía*; Imprenta de D. Francisco de Paula Perojo: Madrid, Spain, 1807.
20. Galet, P. *A Practical Ampelography: Grapevine Identification*; Cornell University Press: Ithaca, NY, USA, 1979.
21. Viala, P.; Vermorel, V. *Traité Général de Viticulture: Ampélographie*; Masson et Cie: Paris, France, 1901.
22. Bodor-Pesti, P.; Taranyi, D.; Deák, T.; Nyitrainé Sárdy, D.Á.; Varga, Z. A Review of Ampelometry: Morphometric Characterization of the Grape (*Vitis* Spp.) Leaf. *Plants* **2023**, *12*, 452. [CrossRef] [PubMed]
23. Chitwood, D.H.; Ranjan, A.; Martinez, C.C.; Headland, L.R.; Thiem, T.; Kumar, R.; Covington, M.F.; Hatcher, T.; Naylor, D.T.; Zimmerman, S.; et al. A Modern Ampelography: A Genetic Basis for Leaf Shape and Venation Patterning in Grape. *Plant Physiol.* **2014**, *164*, 259–272. [CrossRef] [PubMed]
24. Demmings, E.M.; Williams, B.R.; Lee, C.-R.; Barba, P.; Yang, S.; Hwang, C.-F.; Reisch, B.I.; Chitwood, D.H.; Londo, J.P. Quantitative Trait Locus Analysis of Leaf Morphology Indicates Conserved Shape Loci in Grapevine. *Front. Plant Sci.* **2019**, *10*, 1373. [CrossRef] [PubMed]
25. Zecca, G.; Abbott, J.R.; Sun, W.-B.; Spada, A.; Sala, F.; Grassi, F. The Timing and the Mode of Evolution of Wild Grapes (*Vitis*). *Mol. Phylogenet. Evol.* **2012**, *62*, 736–747. [CrossRef] [PubMed]
26. Shi, X.; Cao, S.; Wang, X.; Huang, S.; Wang, Y.; Liu, Z.; Liu, W.; Leng, X.; Peng, Y.; Wang, N.; et al. The Complete Reference Genome for Grapevine (*Vitis vinifera* L.) Genetics and Breeding. *Hortic. Res.* **2023**, *10*, uhad061. [CrossRef] [PubMed]
27. Tympakianakis, S.; Trantas, E.; Avramidou, E.V.; Ververidis, F. *Vitis vinifera* Genotyping Toolbox to Highlight Diversity and Germplasm Identification. *Front. Plant Sci.* **2023**, *14*, 1139647. [CrossRef] [PubMed]
28. Bowers, J.E.; Meredith, C.P. The Parentage of a Classic Wine Grape, Cabernet Sauvignon. *Nat. Genet.* **1997**, *16*, 84–87. [CrossRef] [PubMed]
29. Lacombe, T.; Boursiquot, J.-M.; Laucou, V.; Di Vecchi-Staraz, M.; Péros, J.-P.; This, P. Large-Scale Parentage Analysis in an Extended Set of Grapevine Cultivars (*Vitis vinifera* L.). *Theor. Appl. Genet.* **2013**, *126*, 401–414. [CrossRef] [PubMed]
30. Laucou, V.; Launay, A.; Bacilieri, R.; Lacombe, T.; Adam-Blondon, A.-F.; Bérard, A.; Chauveau, A.; de Andrés, M.T.; Hausmann, L.; Ibáñez, J.; et al. Extended Diversity Analysis of Cultivated Grapevine *Vitis vinifera* with 10K Genome-Wide SNPs. *PLoS ONE* **2018**, *13*, e0192540. [CrossRef] [PubMed]
31. Cervantes, E.; Martín-Gómez, J.J.; Rodríguez-Lorenzo, J.L.; Del Pozo, D.G.; Cabello Sáenz De Santamaría, F.; Muñoz-Organero, G.; Tocino, Á. Seed Morphology in *Vitis* Cultivars Related to Hebén. *AgriEngineering* **2025**, *7*, 62. [CrossRef]
32. Zinelabidine, L.H.; Cunha, J.; Eiras-Dias, J.E.; Cabello, F.; Martínez-Zapater, J.M.; Ibáñez, J. Pedigree Analysis of the Spanish Grapevine Cultivar "Hebén". *VITIS* **2015**, *54*, 81–86. [CrossRef]
33. Zinelabidine, L.H.; Haddioui, A.; Rodríguez, V.; Cabello, F.; Eiras-Dias, J.E.; Martínez-Zapater, J.M.; Ibáñez, J. Identification by SNP Analysis of a Major Role for Cayetana Blanca in the Genetic Network of Iberian Peninsula Grapevine Varieties. *Am. J. Enol. Vitic.* **2012**, *63*, 121–126. [CrossRef]
34. Vavilov, N.I. *Origin and Geography of Cultivated Plants*; Cambridge University Press: Cambridge, UK, 1992.

35. Negrul, A.M. Origin and Classification of Cultivated Grapevines. In *Grapes of the USSR*; Pishchepromizdat: Moscow-Leningrad, Union of Soviet Socialist Republics, 1946; Volume 1.
36. Arroyo-García, R.; Ruiz-García, L.; Bolling, L.; Ocete, R.; López, M.A.; Arnold, C.; Ergul, A.; Söylemezoğlu, G.; Uzun, H.I.; Cabello, F.; et al. Multiple Origins of Cultivated Grapevine (*Vitis vinifera* L. ssp. *sativa*) Based on Chloroplast DNA Polymorphisms. *Mol. Ecol.* **2006**, *15*, 3707–3714. [CrossRef] [PubMed]
37. Badouin, H.; Velt, A.; Gindraud, F.; Flutre, T.; Dumas, V.; Vautrin, S.; Marande, W.; Corbi, J.; Sallet, E.; Ganofsky, J.; et al. The Wild Grape Genome Sequence Provides Insights into the Transition from Dioecy to Hermaphroditism during Grape Domestication. *Genome Biol.* **2020**, *21*, 223. [CrossRef] [PubMed]
38. De Andrés, M.T.; Benito, A.; Pérez-Rivera, G.; Ocete, R.; Lopez, M.A.; Gaforio, L.; Muñoz, G.; Cabello, F.; Martínez-Zapater, J.M.; Arroyo-García, R. Genetic Diversity of Wild Grapevine Populations in Spain and Their Genetic Relationships with Cultivated Grapevines. *Mol. Ecol.* **2012**, *21*, 800–816. [CrossRef] [PubMed]
39. D'Onofrio, C. Introgression Among Cultivated and Wild Grapevine in Tuscany. *Front. Plant Sci.* **2020**, *11*, 202. [CrossRef] [PubMed]
40. Liu, Z.; Wang, N.; Su, Y.; Long, Q.; Peng, Y.; Shangguan, L.; Zhang, F.; Cao, S.; Wang, X.; Ge, M.; et al. Grapevine Pangenome Facilitates Trait Genetics and Genomic Breeding. *Nat. Genet.* **2024**, *56*, 2804–2814. [CrossRef] [PubMed]
41. Xiao, H.; Wang, Y.; Liu, W.; Shi, X.; Huang, S.; Cao, S.; Long, Q.; Wang, X.; Liu, Z.; Xu, X.; et al. Impacts of Reproductive Systems on Grapevine Genome and Breeding. *Nat. Commun.* **2025**, *16*, 2031. [CrossRef] [PubMed]
42. Velt, A.; Frommer, B.; Blanc, S.; Holtgräwe, D.; Duchêne, É.; Dumas, V.; Grimplet, J.; Hugueney, P.; Kim, C.; Lahaye, M.; et al. An Improved Reference of the Grapevine Genome Reasserts the Origin of the PN40024 Highly Homozygous Genotype. *G3 Genes Genomes Genet.* **2023**, *13*, jkad067. [CrossRef] [PubMed]
43. Jaillon, O.; Aury, J.-M.; Noel, B.; Policriti, A.; Clepet, C.; Casagrande, A.; Choisne, N.; Aubourg, S.; Vitulo, N.; Jubin, C.; et al. The Grapevine Genome Sequence Suggests Ancestral Hexaploidization in Major Angiosperm Phyla. *Nature* **2007**, *449*, 463–467. [CrossRef] [PubMed]
44. Velasco, R.; Zharkikh, A.; Troggio, M.; Cartwright, D.A.; Cestaro, A.; Pruss, D.; Pindo, M.; FitzGerald, L.M.; Vezzulli, S.; Reid, J.; et al. A High Quality Draft Consensus Sequence of the Genome of a Heterozygous Grapevine Variety. *PLoS ONE* **2007**, *2*, e1326. [CrossRef] [PubMed]
45. Zhou, Y.; Massonnet, M.; Sanjak, J.S.; Cantu, D.; Gaut, B.S. Evolutionary Genomics of Grape (*Vitis vinifera* ssp. *vinifera*) Domestication. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 11715–11720. [CrossRef] [PubMed]
46. Canaguier, A.; Grimplet, J.; Di Gaspero, G.; Scalabrin, S.; Duchêne, E.; Choisne, N.; Mohellibi, N.; Guichard, C.; Rombauts, S.; Le Clainche, I.; et al. A New Version of the Grapevine Reference Genome Assembly (12X.v2) and of Its Annotation (VCost.V3). *Genom. Data* **2017**, *14*, 56–62. [CrossRef] [PubMed]
47. He, G.-Q.; Jin, H.-Y.; Cheng, Y.-Z.; Yu, Y.-H.; Guo, D.-L. Characterization of Genome-Wide Long Terminal Repeat Retrotransposons Provide Insights into Trait Evolution of Four Grapevine Species. *J. Syst. Evol.* **2023**, *61*, 414–427. [CrossRef]
48. Park, M.; Sarkhosh, A.; Tsolova, V.; El-Sharkawy, I. Horizontal Transfer of LTR Retrotransposons Contributes to the Genome Diversity of *Vitis*. *Int. J. Mol. Sci.* **2021**, *22*, 10446. [CrossRef] [PubMed]
49. Grimplet, J.; Adam-Blondon, A.-F.; Bert, P.-F.; Bitz, O.; Cantu, D.; Davies, C.; Delrot, S.; Pezzotti, M.; Rombauts, S.; Cramer, G.R. The Grapevine Gene Nomenclature System. *BMC Genom.* **2014**, *15*, 1077. [CrossRef] [PubMed]
50. Mejía, N.; Soto, B.; Guerrero, M.; Casanueva, X.; Houel, C.; de los Ángeles Miccono, M.; Ramos, R.; Le Cunff, L.; Boursiquot, J.-M.; Hinrichsen, P.; et al. Molecular, Genetic and Transcriptional Evidence for a Role of *VvAGL11* in Stenospermocarpic Seedlessness in Grapevine. *BMC Plant Biol.* **2011**, *11*, 57. [CrossRef] [PubMed]
51. Bogs, J.; Jaffé, F.W.; Takos, A.M.; Walker, A.R.; Robinson, S.P. The Grapevine Transcription Factor VvMYBPA1 Regulates Proanthocyanidin Synthesis during Fruit Development. *Plant Physiol.* **2007**, *143*, 1347–1361. [CrossRef] [PubMed]
52. Cavallini, E.; Zenoni, S.; Finezzo, L.; Guzzo, F.; Zamboni, A.; Avesani, L.; Tornielli, G.B. Functional Diversification of Grapevine MYB5a and MYB5b in the Control of Flavonoid Biosynthesis in a Petunia Anthocyanin Regulatory Mutant. *Plant Cell Physiol.* **2014**, *55*, 517–534. [CrossRef] [PubMed]
53. Wang, N.; Zheng, Y.; Xin, H.; Fang, L.; Li, S. Comprehensive Analysis of NAC Domain Transcription Factor Gene Family in *Vitis vinifera*. *Plant Cell Rep.* **2013**, *32*, 61–75. [CrossRef] [PubMed]
54. Wan, S.; Li, W.; Zhu, Y.; Liu, Z.; Huang, W.; Zhan, J. Genome-Wide Identification, Characterization and Expression Analysis of the Auxin Response Factor Gene Family in *Vitis vinifera*. *Plant Cell Rep.* **2014**, *33*, 1365–1375. [CrossRef] [PubMed]
55. Ahmad, B.; Zhang, S.; Yao, J.; Rahman, M.U.; Hanif, M.; Zhu, Y.; Wang, X. Genomic Organization of the B3-Domain Transcription Factor Family in Grapevine (*Vitis vinifera* L.) and Expression during Seed Development in Seedless and Seeded Cultivars. *Int. J. Mol. Sci.* **2019**, *20*, 4553. [CrossRef] [PubMed]

56. Jiu, S.; Xu, Y.; Wang, J.; Wang, L.; Wang, S.; Ma, C.; Guan, L.; Abdullah, M.; Zhao, M.; Xu, W.; et al. Genome-Wide Identification, Characterization, and Transcript Analysis of the TCP Transcription Factors in *Vitis vinifera*. *Front. Genet.* **2019**, *10*, 1276. [CrossRef] [PubMed]
57. Chen, J.; Glémin, S.; Lascoux, M. Genetic Diversity and the Efficacy of Purifying Selection across Plant and Animal Species. *Mol. Biol. Evol.* **2017**, *34*, 1417–1428. [CrossRef] [PubMed]
58. Cao, J.; Schneeberger, K.; Ossowski, S.; Günther, T.; Bender, S.; Fitz, J.; Koenig, D.; Lanz, C.; Stegle, O.; Lippert, C.; et al. Whole-Genome Sequencing of Multiple *Arabidopsis thaliana* Populations. *Nat. Genet.* **2011**, *43*, 956–963. [CrossRef] [PubMed]
59. Huang, X.; Kurata, N.; Wei, X.; Wang, Z.-X.; Wang, A.; Zhao, Q.; Zhao, Y.; Liu, K.; Lu, H.; Li, W.; et al. A Map of Rice Genome Variation Reveals the Origin of Cultivated Rice. *Nature* **2012**, *490*, 497–501. [CrossRef] [PubMed]
60. Lam, H.-M.; Xu, X.; Liu, X.; Chen, W.; Yang, G.; Wong, F.-L.; Li, M.-W.; He, W.; Qin, N.; Wang, B.; et al. Resequencing of 31 Wild and Cultivated Soybean Genomes Identifies Patterns of Genetic Diversity and Selection. *Nat. Genet.* **2010**, *42*, 1053–1059. [CrossRef] [PubMed]
61. Magris, G.; Jurman, I.; Fornasiero, A.; Paparelli, E.; Schwope, R.; Marroni, F.; Di Gaspero, G.; Morgante, M. The Genomes of 204 *Vitis vinifera* Accessions Reveal the Origin of European Wine Grapes. *Nat. Commun.* **2021**, *12*, 7240. [CrossRef] [PubMed]
62. Hufford, M.B.; Xu, X.; Van Heerwaarden, J.; Pyhäjärvi, T.; Chia, J.-M.; Cartwright, R.A.; Elshire, R.J.; Glaubitz, J.C.; Guill, K.E.; Kaeppler, S.M.; et al. Comparative Population Genomics of Maize Domestication and Improvement. *Nat. Genet.* **2012**, *44*, 808–811. [CrossRef] [PubMed]
63. Tuskan, G.A.; DiFazio, S.; Jansson, S.; Bohlmann, J.; Grigoriev, I.; Hellsten, U.; Putnam, N.; Ralph, S.; Rombauts, S.; Salamov, A.; et al. The Genome of Black Cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **2006**, *313*, 1596–1604. [CrossRef] [PubMed]
64. Muyle, A.; Zemp, N.; Deschamps, C.; Mousset, S.; Widmer, A.; Marais, G.A.B. Rapid *De Novo* Evolution of X Chromosome Dosage Compensation in *Silene latifolia*, a Plant with Young Sex Chromosomes. *PLoS Biol.* **2012**, *10*, e1001308. [CrossRef] [PubMed]
65. Yu, Y.; Chen, C.; Huang, M.; Yu, Q.; Du, D.; Mattia, M.R.; Gmitter, F.G., Jr. Genetic Diversity and Population Structure Analysis of *Citrus* Germplasm with Single Nucleotide Polymorphism Markers. *J. Am. Soc. Hortic. Sci.* **2018**, *143*, 399–408. [CrossRef]
66. Levadoux, L. Les Populations Sauvages et Cultivées de *Vitis vinifera* L. *Ann. l'Amélior. Plantes* **1956**, *6*, 59–117.
67. Mangafa, M.; Kotsakis, K. A New Method for the Identification of Wild and Cultivated Charred Grape Seeds. *J. Archaeol. Sci.* **1996**, *23*, 409–418. [CrossRef]
68. Obón, C.; Rivera-Obón, D.J.; Valera, J.; Matilla, G.; Alcaraz, F.; Maghradze, D.; Kikvadze, M.; Ocete, C.-A.; Ocete, R.; Nebish, A.; et al. Is There a Domestication Syndrome in *Vitis* (Vitaceae) Seed Morphology? *Genet. Resour. Crop Evol.* **2025**, *72*, 1541–1565. [CrossRef]
69. Terral, J.-F.; Tabard, E.; Bouby, L.; Ivorra, S.; Pastor, T.; Figueiral, I.; Picq, S.; Chevance, J.-B.; Jung, C.; Fabre, L.; et al. Evolution and History of Grapevine (*Vitis vinifera*) under Domestication: New Morphometric Perspectives to Understand Seed Domestication Syndrome and Reveal Origins of Ancient European Cultivars. *Ann. Bot.* **2010**, *105*, 443–455. [CrossRef] [PubMed]
70. Kuhl, F.P.; Giardina, C.R. Elliptic Fourier Features of a Closed Contour. *Comput. Graph. Image Process.* **1982**, *18*, 236–258. [CrossRef]
71. Bonhomme, V.; Picq, S.; Gaucherel, C.; Claude, J. Momocs: Outline Analysis Using R. *J. Stat. Softw.* **2014**, *56*, 1–24. [CrossRef]
72. Martín-Gómez, J.J.; Rodríguez-Lorenzo, J.L.; Espinosa-Roldán, F.E.; Cabello Sáenz de Santamaría, F.; Muñoz-Organero, G.; Tocino, Á.; Cervantes, E. Seed Morphometry Reveals Two Major Groups in Spanish Grapevine Cultivars. *Plants* **2025**, *14*, 1522. [CrossRef] [PubMed]
73. Martín-Gómez, J.J.; Rodríguez-Lorenzo, J.L.; Espinosa-Roldán, F.E.; Cabello Sáenz De Santamaría, F.; Muñoz-Organero, G.; Tocino, Á.; Cervantes, E. Morphometric Analysis Reveals New Data in the History of *Vitis* Cultivars. *Plants* **2025**, *14*, 2481. [CrossRef] [PubMed]
74. Rodríguez-Lorenzo, J.L.; Martín-Gómez, J.J.; Beloqui, Á.A.; de Santamaría, F.C.S.; Organero, G.M.; Tocino, Á.; Cervantes, E. Seed Morphological Changes Associated with the Domestication of *Vitis* Cultivars. *Phyton*, 2026; *in press*. [CrossRef]
75. Cervantes, E.; Tocino, A. Geometric Analysis of *Arabidopsis* Root Apex Reveals a New Aspect of the Ethylene Signal Transduction Pathway in Development. *J. Plant Physiol.* **2005**, *162*, 1038–1045. [CrossRef] [PubMed]
76. Dong, Y.; Duan, S.; Xia, Q.; Liang, Z.; Dong, X.; Margaryan, K.; Musayev, M.; Goryslavets, S.; Zdunić, G.; Bert, P.-F.; et al. Dual Domestications and Origin of Traits in Grapevine Evolution. *Science* **2023**, *379*, 892–901. [CrossRef] [PubMed]
77. Azuma, A.; Kobayashi, S.; Mitani, N.; Shiraishi, M.; Yamada, M.; Ueno, T.; Kono, A.; Yakushiji, H.; Koshita, Y. Genomic and Genetic Analysis of Myb-Related Genes That Regulate Anthocyanin Biosynthesis in Grape Berry Skin. *Theor. Appl. Genet.* **2008**, *117*, 1009–1019. [CrossRef] [PubMed]
78. Kobayashi, S.; Goto-Yamamoto, N.; Hirochika, H. Retrotransposon-Induced Mutations in Grape Skin Color. *Science* **2004**, *304*, 982. [CrossRef] [PubMed]

79. Lijavetzky, D.; Ruiz-García, L.; Cabezas, J.A.; De Andrés, M.T.; Bravo, G.; Ibáñez, A.; Carreño, J.; Cabello, F.; Ibáñez, J.; Martínez-Zapater, J.M. Molecular Genetics of Berry Colour Variation in Table Grape. *Mol. Genet. Genom.* **2006**, *276*, 427–435. [CrossRef] [PubMed]
80. Maestri, S.; Gambino, G.; Lopatriello, G.; Minio, A.; Perrone, I.; Cosentino, E.; Giovannone, B.; Marcolungo, L.; Alfano, M.; Rombauts, S.; et al. 'Nebbiolo' Genome Assembly Allows Surveying the Occurrence and Functional Implications of Genomic Structural Variations in Grapevines (*Vitis vinifera* L.). *BMC Genom.* **2022**, *23*, 159. [CrossRef] [PubMed]
81. Roach, M.J.; Johnson, D.L.; Bohlmann, J.; Van Vuuren, H.J.J.; Jones, S.J.M.; Pretorius, I.S.; Schmidt, S.A.; Borneman, A.R. Population Sequencing Reveals Clonal Diversity and Ancestral Inbreeding in the Grapevine Cultivar Chardonnay. *PLoS Genet.* **2018**, *14*, e1007807. [CrossRef] [PubMed]
82. Frommer, B.; Müllner, S.; Holtgräwe, D.; Viehöver, P.; Huettel, B.; Töpfer, R.; Weisshaar, B.; Zyprian, E. Phased Grapevine Genome Sequence of an Rpv12 Carrier for Biotechnological Exploration of Resistance to *Plasmopara viticola*. *Front. Plant Sci.* **2023**, *14*, 1180982. [CrossRef] [PubMed]
83. Minio, A.; Cochetel, N.; Massonnet, M.; Figueroa-Balderas, R.; Cantu, D. HiFi Chromosome-Scale Diploid Assemblies of the Grape Rootstocks 110R, Kober 5BB, and 101–14 Mgt. *Sci. Data* **2022**, *9*, 660. [CrossRef] [PubMed]
84. Long, Q.; Cao, S.; Huang, G.; Wang, X.; Liu, Z.; Liu, W.; Wang, Y.; Xiao, H.; Peng, Y.; Zhou, Y. Population Comparative Genomics Discovers Gene Gain and Loss during Grapevine Domestication. *Plant Physiol.* **2024**, *195*, 1401–1413. [CrossRef] [PubMed]
85. Li, B.; Gschwend, A.R. *Vitis labrusca* Genome Assembly Reveals Diversification between Wild and Cultivated Grapevine Genomes. *Front. Plant Sci.* **2023**, *14*, 1234130. [CrossRef] [PubMed]
86. Wang, Y.; Xin, H.; Fan, P.; Zhang, J.; Liu, Y.; Dong, Y.; Wang, Z.; Yang, Y.; Zhang, Q.; Ming, R.; et al. The Genome of Shanputao (*Vitis amurensis*) Provides a New Insight into Cold Tolerance of Grapevine. *Plant J.* **2021**, *105*, 1495–1506. [CrossRef] [PubMed]
87. Luo, Y.; Liu, Z.; Jin, Z.; Li, P.; Tan, X.; Cao, S.; Wang, X.; Liu, Z.; Shi, X.; Huang, S.; et al. Phased T2T Genome Assemblies Facilitate the Mining of Disease-Resistance Genes in *Vitis davidii*. *Hortic. Res.* **2025**, *12*, uhae306. [CrossRef] [PubMed]
88. Girollet, N.; Rubio, B.; Lopez-Roques, C.; Valière, S.; Ollat, N.; Bert, P.-F. *De Novo* Phased Assembly of the *Vitis riparia* Grape Genome. *Sci. Data* **2019**, *6*, 127. [CrossRef] [PubMed]
89. Cochetel, N.; Minio, A.; Guarracino, A.; Garcia, J.F.; Figueroa-Balderas, R.; Massonnet, M.; Kasuga, T.; Londo, J.P.; Garrison, E.; Gaut, B.S.; et al. A Super-Pangenome of the North American Wild Grape Species. *Genome Biol.* **2023**, *24*, 290. [CrossRef] [PubMed]
90. Wang, N.; Chen, P.; Xu, Y.; Guo, L.; Li, X.; Yi, H.; Larkin, R.M.; Zhou, Y.; Deng, X.; Xu, Q. Phased Genomics Reveals Hidden Somatic Mutations and Provides Insight into Fruit Development in Sweet Orange. *Hortic. Res.* **2024**, *11*, uhad268. [CrossRef] [PubMed]
91. Sichel, V.; Sarah, G.; Girollet, N.; Laucou, V.; Roux, C.; Bert, P.-F.; Le Cunff, L.; This, P.; Lacombe, T. Chimeras in Merlot Grapevine Revealed by Phased Assembly. *BMC Genom.* **2023**, *24*, 396. [CrossRef] [PubMed]
92. Myles, S.; Boyko, A.R.; Owens, C.L.; Brown, P.J.; Grassi, F.; Aradhya, M.K.; Prins, B.; Reynolds, A.; Chia, J.-M.; Ware, D.; et al. Genetic Structure and Domestication History of the Grape. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 3530–3535. [CrossRef] [PubMed]
93. Arnold, C.; Schnitzler, A.; Douard, A.; Peter, R.; Gillet, F. Is There a Future for Wild Grapevine (*Vitis vinifera* subsp. *silvestris*) in the Rhine Valley? *Biodivers. Conserv.* **2005**, *14*, 1507–1523. [CrossRef]
94. Maraš, V.; Tello, J.; Gazivoda, A.; Mugoša, M.; Perišić, M.; Raičević, J.; Štajner, N.; Ocete, R.; Božović, V.; Popović, T.; et al. Population Genetic Analysis in Old Montenegrin Vineyards Reveals Ancient Ways Currently Active to Generate Diversity in *Vitis vinifera*. *Sci. Rep.* **2020**, *10*, 15000. [CrossRef] [PubMed]
95. Wang, X.; Sun, J.; Wang, J.; Zhang, X.; Chen, S.; Jin, J.; Zhang, X.; Khan, F.S.; Wang, K.; Mei, J.; et al. Origin and Evolution of Grapevine Genomes. *bioRxiv* **2026**. [CrossRef]
96. Zhou, Y.; Minio, A.; Massonnet, M.; Solares, E.; Lv, Y.; Beridze, T.; Cantu, D.; Gaut, B.S. The Population Genetics of Structural Variants in Grapevine Domestication. *Nat. Plants* **2019**, *5*, 965–979. [CrossRef] [PubMed]
97. Mattivi, F.; Guzzon, R.; Vrhovsek, U.; Stefanini, M.; Velasco, R. Metabolite Profiling of Grape: Flavonols and Anthocyanins. *J. Agric. Food Chem.* **2006**, *54*, 7692–7702. [CrossRef] [PubMed]
98. Ribéreau-Gayon, P. *Recherches sur les Anthocyanes des Végétaux: Application Au Genre Vitis*; Librairie Générale de l'Enseignement: Paris, France, 1959.
99. Chitwood, D.H.; Torres-Lomas, E.; Hadi, E.S.; Peterson, W.L.G.; Fischer, M.F.; Rogers, S.E.; He, C.; Acierno, M.G.F.; Azumaya, S.; Benjamin, S.W.; et al. A High-resolution Model of the Grapevine Leaf Morphospace Predicts Synthetic Leaves. *Plants People Planet* **2025**, *7*, 133–146. [CrossRef]
100. Lukšić, K.; Zdunić, G.; Mucalo, A.; Marinov, L.; Ranković-Vasić, Z.; Ivanović, J.; Nikolić, D. Microstructure of Croatian Wild Grapevine (*Vitis vinifera* subsp. *sylvestris* Gmel Hegi) Pollen Grains Revealed by Scanning Electron Microscopy. *Plants* **2022**, *11*, 1479. [CrossRef] [PubMed]

101. Martín-Gómez, J.J.; Cunha, J.; Rodríguez-Lorenzo, J.L.; Anocibar Beloqui, Á.; De Santa María, F.C.S.; Muñoz-Organero, G.; Tocino, Á.; Cervantes, E. Biodiversity for Sustainable Viticulture: Seed Morphometry in Portuguese Cultivars of *Vitis vinifera* L. *Horticulturae* **2026**, *12*, 634. [CrossRef]
102. Biselli, C.; Fricano, A.; Crespan, M.; Giacosa, S.; Garavelloni, S.; Valentini, F.M.; Epifani, A.M.; Espinosa-Roldán, F.E.; Rolle, L. A Comprehensive SPET-Based Linkage Mapping for Berry Texture and Defective Seed Development in Grapevine. *Front. Plant Sci.* **2026**, *17*, 1804241. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.