

Review

Genetic Diversity in *Vitis vinifera* L. Beyond the Reference Genome: Towards a Pangenomic Framework for Representation, Adaptation and Breeding

Francesca Fort ^{1,*} , Leonor Deis ^{1,2}, Qiying Lin-Yang ¹ , Joan Miquel Canals ¹  and Fernando Zamora ¹ 

¹ Grupo de Tecnología Enológica (TECNENOL), Department of Biochemistry and Biotechnology, Faculty of Oenology, Rovira i Virgili University, Sescelades Campus, C/ Marcel·lí Domingo, 1, E-43007 Tarragona, Spain; leonor.deis@urv.cat (L.D.); qiying.lin@urv.cat (Q.L.-Y.); jmcanals@urv.cat (J.M.C.); fernando.zamora@urv.cat (F.Z.)

² Plant Physiology, Faculty of Agricultural Sciences, National University of Cuyo, Mendoza M5528AHB, Argentina

* Correspondence: mariafrancesca.fort@urv.cat

Abstract

The growing availability of genomic resources is changing how genetic diversity is studied in *Vitis vinifera* L. At the same time, it has become increasingly clear that a single reference genome cannot fully represent the complexity of a species characterised by high heterozygosity, clonal propagation and a long history of diversification. Recent grapevine pangenomes, super-pangenomes and graph-based resources have revealed forms of variation that are often overlooked in conventional reference-based analyses, including structural variants and gene presence–absence variation. Rather than providing another inventory of available datasets, this review examines how continued reliance on a single reference genome may influence the interpretation of grapevine diversity and what can be gained from a broader pangenomic perspective. Drawing on recent studies in grapevine and other crops, we discuss how these approaches are beginning to improve the representation of genetic diversity, uncover biologically relevant variation and strengthen links between genomic information and adaptive traits. We also examine the challenges that still limit their practical use, particularly the integration of genomic resources with functional studies and breeding programmes. In the end, the value of pangenomics will probably depend not only on generating additional genomic resources, but also on how effectively these can be translated into tools that support grapevine conservation, climate adaptation and varietal improvement.

Keywords: pangenomics; *Vitis vinifera* L.; genetic diversity; structural variation; gene presence/absence variation; representation bias; phenotype integration; grapevine breeding; climate adaptation



Academic Editors: Alessandra Zombardo and Clizia Villano

Received: 11 May 2026

Revised: 14 June 2026

Accepted: 17 June 2026

Published: 21 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The study of genetic diversity in *Vitis vinifera* L. has traditionally relied on reference genomes. Since the first assemblies became available, they have provided a common framework for organising variation, comparing plant material across studies, and reconstructing genetic relationships [1–4].

A closer analysis indicates, however, that this model begins to show its limitations not only in terms of resolution, but also in how genetic diversity is framed. Grapevine is shaped

by vegetative propagation, and the accumulation of somatic mutations is not marginal, as illustrated by clonal-level variability observed in traditional vineyard systems [5]. Heterozygosity is also high in many genotypes. These factors favour the accumulation of structural variation over time [6], which may complicate the interpretation of genetic diversity when different genotypes are analysed against a single reference genome [7].

As genomic resources continue to expand, these limitations become more apparent. Some genomic regions are present in particular genotypes but absent in others, and this variation may involve structural rearrangements or differences in gene content that are poorly represented in a single-reference framework [8,9]. Recent grapevine assemblies support this view, showing that gene presence/absence variation is not a marginal component of diversity in this species [10].

In this context, the concept of the pangenome becomes particularly relevant. It is based on the premise that the complete gene repertoire of a species is not identical across all genotypes [11,12]. Consequently, the interpretation of genetic diversity extends beyond variation detected relative to a single reference genome.

The increasing availability of high-quality genome assemblies has made it easier to move beyond the limitations of relying on a single reference genome. As a consequence, pangenomes, graph-based representations and, more recently, super-pangenomes are increasingly being incorporated into grapevine genomics. Their value lies not only in expanding the range of detectable variants, but also in making visible forms of diversity that often remain overlooked in conventional reference-based analyses, including structural variants and differences in gene presence or absence. This broader representation also provides new opportunities to connect genomic diversity with functional studies and association analyses [10,13,14]. This is particularly important for complex traits, where part of the underlying variation may extend beyond SNPs or small indels and therefore remain only partially captured when analyses depend on a single reference genome.

Progress is evident, but the framework is not yet consolidated [10]. This leads to a broader question about the extent to which continued reliance on reference genomes may shape the interpretation of genetic diversity in grapevine, and how much of that diversity is still not captured by the frameworks currently used to describe and compare genotypes.

This work addresses that question by examining the current status of pangenomics in grapevine, in comparison with other crops where these approaches are more developed, and by assessing how the reference genome continues to condition the interpretation of diversity in this species.

Recent reviews have summarised the expansion of grapevine genomic resources and the emergence of pangenomic datasets [15]. The aim of the present review is therefore not to provide another inventory of available resources but to examine what these resources change in the interpretation of grapevine diversity. More specifically, we focus on three related questions: how pangenomic frameworks improve the representation of genetic diversity, how they may help identify functional variation that remains poorly captured in reference-based analyses, and how this information could be translated into conservation, adaptation and breeding strategies. In this sense, this review is intended as a critical synthesis of the current transition from genome description towards biological and practical use in grapevine.

2. Literature Search Strategy

The literature considered in this review was gathered through searches in Scopus, Web of Science and Google Scholar, covering publications available up to the time of manuscript revision. The search focused mainly on studies published from the appearance of the first grapevine reference genomes onwards, with particular attention to recent work

on grapevine pangenomics, super-pangenomes, graph-based genome representations, structural variation and gene presence/absence variation.

The searches combined terms such as grapevine pangenome, *Vitis* pangenomics, *Vitis* super-pangenome, graph pangenome, structural variation, presence/absence variation, reference bias, grapevine genomics and crop pangenome. The wording was not identical in all databases since some searches were refined after key papers had been identified. References cited in these papers were also checked when they helped to locate relevant studies that were not retrieved directly in the first searches.

Studies were included when they contributed directly to one of the main aims of this review: understanding how pangenomic approaches affect the representation and interpretation of genetic diversity in grapevine. Priority was given to peer-reviewed articles dealing with grapevine genome assemblies, pangenomic or super-pangenomic resources, structural variation, graph-based approaches, phenotype integration or breeding-related applications. Studies from other crops were included only when they provided a clear comparative point, for example, by showing how pangenomic resources have been connected to trait discovery, functional genomics or breeding use.

More general plant genomics papers, studies with only indirect relevance to pangenomics, and papers that did not provide a useful methodological, biological or applied comparison were not considered in detail. Because the aim was to provide a critical narrative synthesis rather than a systematic review, no formal PRISMA-type screening was applied.

3. What Is a Pangenome? Conceptual Framework and Relevance in Grapevine

The concept of the pangenome is not new. Its use in plants, however, is relatively recent. In simple terms, it refers to the full set of genes that can be found within a species, those shared by all individuals, and those present only in some of them [11,12]. The idea comes from microbial genomics, but its adoption in plants has expanded as more genomes have become available and assemblies have improved [12,16]. The definition is straightforward. What follows from it is less so.

The pangenome is commonly described as consisting of two main components. The core genome is generally understood as the set of genes shared across analysed genotypes, while the variable fraction includes genes that are not found in all individuals, although in practice this distinction depends strongly on sampling. This second fraction can vary substantially among individuals and is often associated with adaptation or specific agronomic traits [12,16–18]. Some studies further distinguish private genes, which are restricted to a single accession and represent the most exclusive fraction of the variable genome. The distinction seems clear on paper. In practice, it depends strongly on how many, and which, genotypes are considered.

One of the key consequences of this framework is the explicit inclusion of gene presence/absence variation (PAV) (Table 1). This type of variation is largely invisible in analyses based on a single reference genome, yet it has been shown to be a significant component of genetic diversity in several crops [16,19,20]. Not all individuals share the same gene repertoire. This affects the way variation is interpreted, particularly when linking genomic differences to phenotype.

Table 1. Main categories of genetic variation considered in pangenomic analyses.

Variant Type	Definition	Example of Relevance in Grapevine
SNP	Single nucleotide polymorphism affecting one nucleotide position	Widely used for varietal identification and population genetics
Indel	Small insertion or deletion of nucleotides	Can alter coding sequences and gene regulation
SV	Large genomic rearrangements including insertions, deletions, inversions or translocations	Frequently associated with clonal variation and genome structural diversity
CNV	Variation in the number of copies of genomic regions or genes	May influence gene dosage and quantitative traits
TE	Transposable element insertion or movement within the genome	Can modify gene expression and generate structural variation
PAV	Presence or absence of genes among genotypes	Important source of diversity captured by pangenomic approaches

In grapevine, this is not just a theoretical possibility. Recent assemblies already show that certain genomic regions are consistently present in some cultivars while absent in others, sometimes involving genes with potential functional relevance [7,10]. These differences are not always captured when variation is forced onto a single reference sequence. They tend to appear indirectly, if at all, which makes their interpretation less straightforward. And, in some cases, they simply remain outside the analysis.

From this perspective, the pangenome is more than a simple extension of the reference genome. It reflects a gradual shift in how diversity is conceptualised. For this reason, it has been proposed as a new reference framework in plant genomics [16,21,22]. The classical model assumes a single sequence onto which variation is projected. The pangenomic approach starts from a different premise: that one sequence is not sufficient.

Early plant pangenomes were built by comparing multiple complete genomes and distinguishing conserved regions from more variable ones [12,21]. This approach has evolved. Graph-based representations now allow genomic diversity to be integrated more flexibly, without forcing alignment onto a single linear reference. This seems particularly relevant for structural variation, which is often difficult to capture under conventional frameworks.

In grapevine, this framework is especially pertinent. It is a species characterised by high intra-specific diversity, shaped by clonal propagation and the accumulation of somatic mutations, where it is reasonable to assume that part of the variation is not fully represented in a single reference genome [6,7]. The value of the pangenome lies, therefore, not only in expanding the catalogue of variants, but in changing how that diversity is interpreted.

In practice, the reference genome is not discarded. It remains useful as a coordinate system and as a point of comparison. The pangenome becomes relevant when the biological question requires variation that cannot be properly represented by a single linear sequence, especially structural variants and gene presence/absence differences (Figure 1).

In practice, these conceptual categories are useful as a framework, but their interpretation is often shaped by methodological choices, which makes direct comparisons between studies less straightforward than they may initially appear.

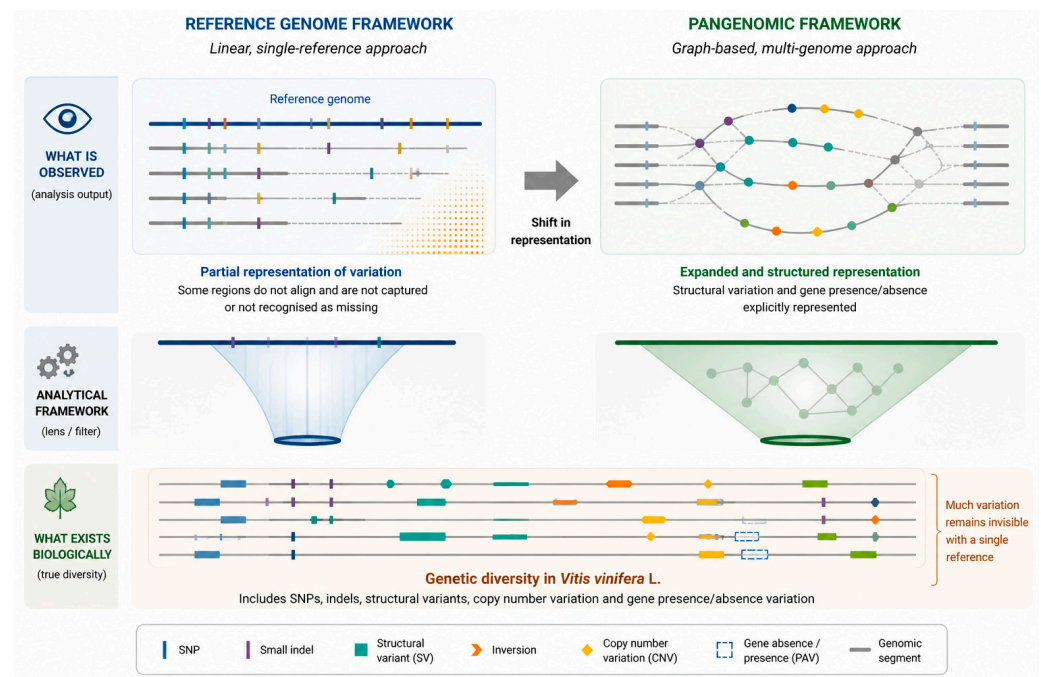


Figure 1. Analytical frameworks and the visibility of genetic diversity in *Vitis vinifera* L. Conceptual diagram illustrating how reference genome–based analyses capture only a fraction of genetic variation, whereas a pangenomic framework enables a more complete and structured representation, including structural variation and gene presence/absence. Conceptual illustration generated with the assistance of AI-based image tools and subsequently edited by the authors.

4. Current Status of Pangenomics in Grapevine

Pangenomics in grapevine has advanced substantially in recent years, although it still cannot be considered a fully consolidated framework. Several studies already point to its potential. At the same time, the level of integration reached in other crops, where pangenomes are used more routinely in functional analyses, association studies or breeding programmes, has not yet been achieved.

The development of the field has also been uneven. Even before formal grapevine pangenomes were available, different studies had already suggested that part of the genetic variation was probably being missed by conventional approaches. Structural variation, clonal differences and changes in gene content were already pointing in that direction, suggesting that part of the diversity being described was not simply missing, but structurally excluded from the analytical framework itself [7–9]. The need for more integrative frameworks was visible early on. At that stage, however, this need had not yet been explicitly framed in pangenomic terms [15].

The transition from conceptual discussions to operational pangenomic resources has accelerated during the last few years (Table 2). One of the most significant developments has been the construction of Grapepan, a grapevine pangenome integrating multiple cultivated genomes and providing a broader representation of gene-content variation than was previously available [10]. Grapepan was built from 18 newly generated phased telomere-to-telomere assemblies and 11 previously published assemblies and was then used to construct a variation map including 9,105,787 short variants and 236,449 structural variants from resequencing data of 466 grapevine cultivars. This began to provide a more concrete basis for integrating structural variation into trait genetics and genomic prediction. At the same time, the North American wild-grape super-pangenome expanded the scope beyond cultivated *Vitis vinifera* L. by incorporating wild *Vitis* species, revealing additional genomic diversity and highlighting previously underrepresented sources of variation [13].

Together, these resources illustrate a transition from proof-of-concept studies towards frameworks that are beginning to connect structural variation, gene-content diversity and trait-associated variation across cultivated and wild *Vitis* germplasm.

Table 2. Major grapevine pangenomic resources and their biological scope, principal contributions and current limitations.

Resource	Scale and Biological Scope	Main Contribution	Current Limitation
Grapepan (Liu et al., 2024) [10]	29 genome assemblies, including 18 newly generated phased telomere-to-telomere assemblies and 11 previously published assemblies	Identification of variable gene fraction and integration of structural variants into trait genetics and genomic prediction	Incomplete representation of full <i>Vitis</i> diversity, especially wild relatives
Cochetel et al., 2023 [13]	Super-pangenome across 11 North American wild <i>Vitis</i> species interspecific genomic diversity	Construction of a super-pangenome capturing previously underrepresented wild <i>Vitis</i> diversity	Limited linkage to phenotypic traits and functional validation
Guo et al., 2025 [14]	Super-pangenome across multiple <i>Vitis</i> species integrating wild and cultivated diversity	Identification and mapping of downy mildew resistance <i>loci</i> using a <i>Vitis</i> super-pangenome	Broader phenotypic applications remain largely unexplored
Liu et al., 2025 [23]	Integration of pangenomics and single-cell transcriptomics to study the continuous bearing trait	Identification of the genetic basis of continuous bearing through combined genomic and cell-type expression analysis	Broader implementation in breeding and germplasm studies remains to be explored

With the availability of higher-quality genome assemblies, it has become possible to move a step further. Pangenomes in *Vitis vinifera* L. have enabled the identification of a variable genomic fraction with a substantial contribution, and have begun to link this variation to complex agronomic traits [10,23]. This changes the analytical scale, not completely, but enough to matter. Genetic diversity is no longer reduced to single-nucleotide variation; differences in gene content and structural organisation become part of the picture.

Even so, representativeness remains a major limitation. Current grapevine pangenomic resources are still based on relatively limited numbers of assembled genomes. Grapepan, for example, integrates 29 assemblies, while the North American wild-grape super-pangenome was built from newly assembled wild-grape genomes and related *Vitis* diversity. These scales are useful, but still modest compared with crops where pangenomic studies are increasingly supported by much larger resequencing panels and more systematic phenotype integration [12,16,21]. In a species with such high intra-specific diversity, this is not a secondary issue. If anything, it sits at the centre of the problem.

Super-pangenomes have recently opened a complementary line of work by integrating cultivated varieties together with wild species of the genus *Vitis*. These approaches reveal additional layers of genetic variation, including regions absent from commonly used reference genomes [13,14,18,24]. In some cases, this variation has already been associated with agronomically relevant traits, such as disease resistance, reinforcing the applied potential of the approach [14]. Still, these links remain limited. They are starting to appear, but they do not yet define the field.

The connection between genomic variation and phenotype remains, in general terms, weak. In other crops, pangenomes are more systematically integrated into association studies and into the identification of variants relevant for breeding [17,20,25]. In grapevine, similar approaches are beginning to appear, but they are still scattered. Some recent studies combine pangenomics with transcriptomics or functional analyses. That is clearly a step forward. It is not yet a stable framework, though. It may still be too early to consider that integration established [23].

The biology of grapevine adds another layer of complexity. High intra-specific diversity, together with the clonal history of many cultivars, affects both genome assembly and

direct comparison between genotypes [6,7]. This is further complicated by reproductive systems and by the evolutionary history of the genus, which may influence how genetic variation is structured [26]. These factors do not prevent the use of pangenomics, but they do shape how it can be applied.

There is also the question of scale. Although the number of grapevine genomes has increased in recent years, it remains lower than in crops where pangenomics is more developed, such as rice, wheat, tomato or sorghum [17,19,20,25]. This difference constrains both the definition of the pangenome and its use in broader comparative analyses.

Overall, pangenomics in grapevine is not so much at an early stage as at an incomplete one. There is already solid evidence of its relevance, but not yet a sufficiently integrated framework. The use of the reference genome as the sole analytical model is becoming increasingly difficult to justify. What is still missing is a stable connection between genomes, structural variants, phenotypes and breeding applications [16,18,21].

At this point, the main issue does not seem to be simply the generation of more genomes. The real difficulty lies in how those genomes are integrated and interpreted. The pangenomic resources currently available differ in scale, scope and biological focus, yet they converge on a similar message: several recent pangenomic resources consistently indicate that structural variation and gene-content variation account for an important component of grapevine diversity, although their exact contribution remains difficult to quantify because current resources differ substantially in scope and sampling. Under these conditions, the challenge is no longer to demonstrate that additional diversity exists, but to understand how that diversity influences biological interpretation and practical applications. More data will undoubtedly be useful, but recent developments in pangenomics and genome evolution suggest that integration, more than data production alone, will ultimately determine whether pangenomics becomes a genuinely operational framework for grapevine research [22].

Cross-Study Synthesis and Current Limitations

Although the grapevine pangenomic resources currently available differ in their biological scope, the overall picture that emerges is relatively consistent. Structural variation and gene presence/absence variation repeatedly appear as important components of diversity, suggesting that part of the genomic variability in *Vitis* cannot be adequately described through single-reference approaches alone. At the same time, direct comparisons between studies remain difficult because the available resources differ substantially in sampling strategy, taxonomic coverage and analytical methodology.

Another recurring limitation concerns representativeness. Current grapevine pangenomes are based on a limited number of assembled genomes when compared with several major crop systems. This does not diminish their value, but it does influence the stability of estimates related to core and variable genomic fractions and may affect how broadly the resulting conclusions can be extrapolated across the species.

A further challenge lies in the connection between genomic variation and phenotype. Recent studies have begun to associate pangenomic variation with agronomically relevant traits, including disease resistance and developmental characteristics, but these examples remain relatively scarce. In most cases, functional validation and phenotype integration still lag behind resource generation.

Taken together, these observations suggest that the principal challenge is no longer the demonstration that additional genomic diversity exists. Rather, it is the development of analytical frameworks capable of integrating structural variation, gene-content variation and phenotypic information into a coherent biological interpretation. The future value of grapevine pangenomics will depend largely on progress in this direction.

5. From Pangenomics in Major Crops to Its Application in Grapevine: A Shift in Scale

Looking at other crops helps to place grapevine in a broader context. Even so, it is worth noting that not all of these advances are directly transferable. In some cases, the progress observed in other crops depends on specific contexts, such as population size, breeding structure or data availability, that do not always have a clear equivalent in grapevine. This does not invalidate the comparison, but it does call for a more cautious interpretation. Table 3 brings these differences together in a more structured way, making it easier to compare how pangenomic approaches have developed across systems.

Table 3. Pangenomic insights from major crops and their implications for the interpretation of genetic diversity in *Vitis vinifera* L.

Crop	Pangenomic Approach	Key Biological Insight	Link to Traits/Function	Transferability to Grapevine
Rice	Population-scale pangenomes distinguishing shared and variable gene content	A substantial fraction of gene content is not consistently present across accessions, and becomes visible only when multiple genomes are analysed together	Stress response, adaptation, yield-related <i>loci</i>	Transferable: broad representation improves detection of hidden diversity. Limitation: annual crop with stronger breeding structure than grapevine
Wheat	Large-scale assemblies integrating multiple genomes, with explicit treatment of structural variation	Structural variation represents a major component of diversity that often extends beyond the limits of linear genome alignments	Agronomic traits frequently associated with gene presence/absence and rearrangements	Transferable: SV must be explicitly represented. Limitation: polyploid genome context differs from grapevine
Tomato	Pangenome combining cultivated material with wild relatives	Rare alleles and introgressions can disproportionately shape complex phenotypes	Fruit quality, flavour, domestication-related traits	Transferable: wild relatives and introgressed regions can reveal trait-associated variation. Limitation: breeding history and reproductive biology differ
Cucumber	Graph-based pangenome integrating breeding populations and structural variation	Structural variation changes dynamically during breeding and contributes substantially to genomic differentiation	Breeding-associated genomic regions and structural variants linked to crop improvement	Transferable: graph approaches improve SV discovery during breeding. Limitation: structured breeding populations are less comparable to grapevine diversity
Sorghum	Global pangenome reference integrating diverse cultivated germplasm	Important components of trait-associated variation become visible only when diversity is represented beyond a single reference genome	Enhanced discovery of <i>loci</i> associated with adaptation and agronomic traits	Transferable: global pangenome resources improve trait discovery. Limitation: grapevine lacks equivalent phenotype-integrated panels
Grapevine	Emerging pangenomic resources, still limited in integration	Presence/absence variation and structural differences have been reported, although often in a fragmented way	Links between structural variation, stress response and adaptation are beginning to emerge	Main need: move from resource generation to phenotype integration, validation and breeding use

What emerges from this comparison is not simply a difference in the amount of available data, but in how these data are organised and interpreted across systems. Interestingly, the transition from data generation to biological application appears to have followed a similar trajectory in most crops. The main differences concern how rapidly this transition has occurred rather than the direction in which it has progressed.

The aim is not to replicate these models directly, but rather to identify the factors that have enabled pangenomics to evolve from a descriptive framework into a practical tool for understanding genetic diversity and supporting breeding programmes [16,21,27].

This distinction is particularly important when interpreting how pangenomics becomes operational (Figure 2).

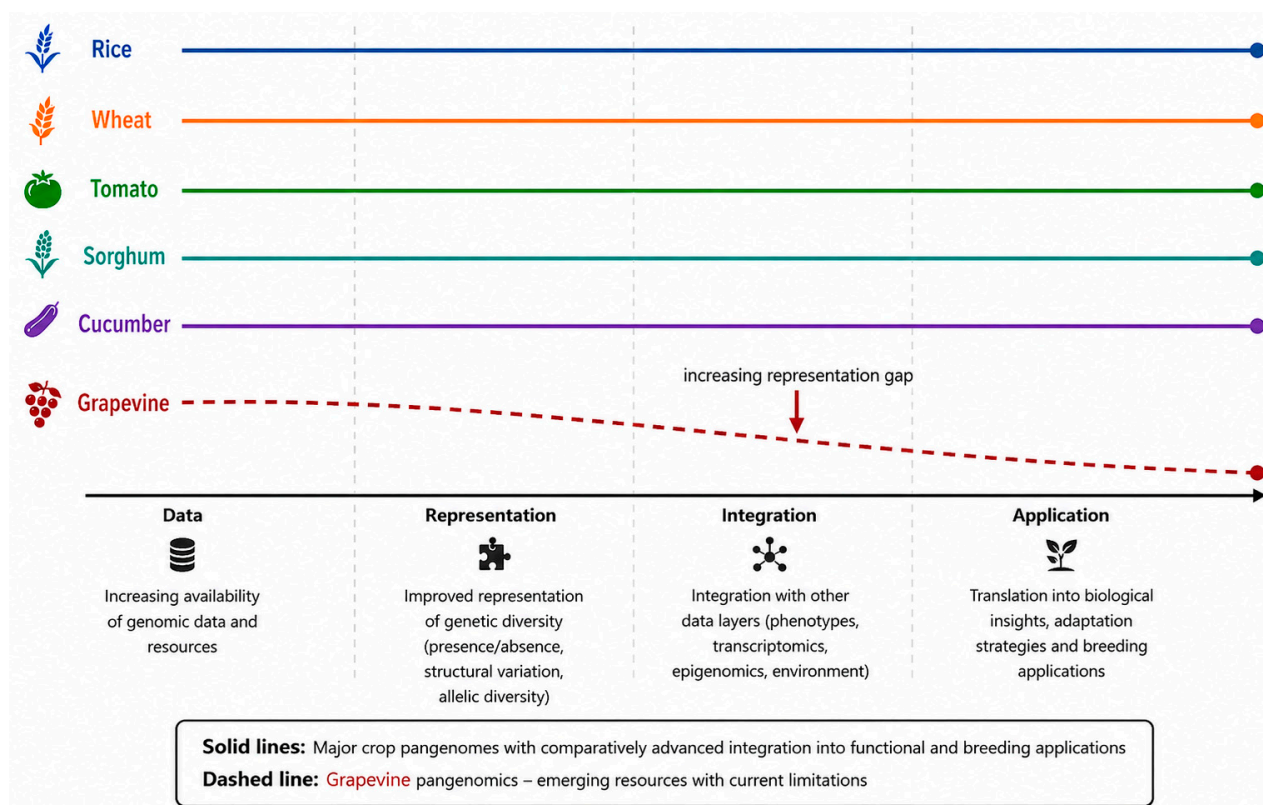


Figure 2. Relative progression of pangenomic frameworks from data generation to practical application across major crops and grapevine. Rice, wheat, tomato, cucumber and sorghum illustrate systems in which pangenomic resources have moved beyond genome description and are increasingly used in functional analyses, trait discovery and breeding. Grapevine shows a similar pattern, although the connection between genomic resources and practical applications remains more limited. The figure offers a conceptual comparison across crop systems. Conceptual illustration generated with the assistance of AI-based image tools and subsequently edited by the authors.

In rice, pangenome analyses showed that genomic diversity was broader than initially inferred from a single reference genome. These studies highlighted the presence of genes and genomic regions not shared across all genotypes, with direct implications for how crop diversity is interpreted [20]. More recently, super-pangenomes have extended this perspective by integrating diversity from both cultivated varieties and wild relatives, uncovering fractions of variation relevant for adaptation and breeding [28,29]. For grapevine, the comparison is informative for a similar reason: in species with a broad genetic base, a single reference may provide order, but not completeness. It provides a useful framework, but not a complete one. This limitation is not resolved simply by increasing the amount of data.

In wheat, the value of the pangenome has been linked mainly to genome complexity and to the need to capture variation absent from early reference sequences. Initial pangenomes already showed the magnitude of the variable fraction, while more recent work has integrated global diversity, including modern varieties, landraces and wild materials, and connected it with specific agronomic traits [19,30]. The genomic context is, of course, different. Still, the underlying message for grapevine is similar, even if the biological details are not.

Tomato provides another useful example. In this crop, pangenomics enabled the identification of rare genes and alleles associated with traits of interest, especially those

related to fruit quality [17]. The introduction of graph-based pangenomes has reinforced this approach by improving the detection of structural variation and previously inaccessible alleles, and by facilitating their integration into association studies and breeding programmes [31]. This is particularly relevant for grapevine, where important traits, berry aroma composition, responses to water or heat stress, or disease resistance, may depend not only on single variants, but also on differences in gene content or structural variation that remain only partially explored [7,9,10]. Not all of that variation is easy to capture with current approaches. Some of it, quite simply, does not align well—and, in some cases, is not recognised as missing in the first place.

Graph-based pangenomes have strengthened this tendency in other crops as well. In cucumber, they have enabled a more detailed analysis of structural variation dynamics during breeding processes [32]. In sorghum, a pangenome reference has improved the identification of variants associated with agronomic traits at a global scale [25]. Methodologically, these advances rely on specific strategies for constructing, analysing and visualising graph-based pangenomes, which allow complex and heterogeneous genomes to be handled more efficiently [33]. Taken together, these examples suggest something relatively consistent across systems: structural and gene-content variation tends to improve the identification of functional variants. It also makes their use in applied analyses more realistic.

Looking across these examples, what stands out is not so much the amount of new variation that has been discovered in each crop, but the fact that similar conclusions keep appearing despite the obvious biological differences among them. Rice, wheat, tomato or sorghum have followed very different evolutionary and breeding histories. Even so, pangenomic studies repeatedly suggest that part of the diversity associated with adaptation, agronomic performance or complex phenotypes remains difficult to capture when analyses rely on a single reference genome. The details change from one species to another, sometimes quite substantially, but the general picture is remarkably consistent. In most cases, the real advance has not come from generating more data alone. It has come from making previously overlooked forms of variation visible and from connecting them to phenotype in ways that were difficult to achieve before. A particularly relevant point is that, across these crops, the practical value of pangenomics did not emerge when new genomes became available, but when those genomes could be connected to traits, biological functions and breeding objectives. In that sense, representation was only the first step.

From these examples, three points seem especially relevant for grapevine. The first is the need for a sufficiently large and diverse set of genotypes. Without broad representation, the distinction between the core genome and the variable fraction depends heavily on sampling and may become unstable [12,16,27]. In grapevine, this is particularly important because cultivated diversity coexists with a wide pool of wild diversity, while many traditional or local varieties remain underrepresented. That imbalance matters more than it may appear.

The second point is the connection between pangenomic variation and phenotype. In crops where pangenomics is more advanced, the decisive step has not only been the construction of variant catalogues. It has been their connection to agronomic traits through association studies, transcriptomic integration or other functional approaches [17,25,34]. In grapevine, there are already initial efforts in this direction, but they are not yet part of standard practice [23]. The gap is not conceptual; it is operational. This point marks a stage at which progress often slows down.

The third point concerns practical utility. In other crops, pangenomes are increasingly used as reference resources for breeding, candidate gene identification and the interpretation of complex traits, and they are progressively being incorporated into standard breeding workflows [16,21,35]. In grapevine, the potential is clear, but the development is slower. The

main bottleneck is therefore operational: genomic resources need to become comparable, reusable and connected to phenotype.

At the same time, grapevine is not simply another crop following the same trajectory with some delay. In rice, wheat, tomato, cucumber or sorghum, the development of pangenomics has often been linked to structured breeding programmes and large-scale datasets. In grapevine, the biological context introduces a different kind of complexity, associated with the historical diversity of cultivars, the persistence of local materials and the contribution of wild *Vitis* species [15,26,36]. This does not weaken the relevance of pangenomics. If anything, it reinforces it, although not necessarily in the same way.

Ultimately, the comparison with other crops suggests that the most important step is not the construction of increasingly larger pangenomes. Most species have already shown that additional genomes continue to reveal new variation. The real challenge begins afterwards, when that variation must be linked to phenotype, adaptation and practical use. Grapevine appears to be approaching that stage. Whether pangenomics becomes a routinely applied framework will probably depend less on the amount of available genomic information than on the ability to translate that information into biological understanding [10,14,23,35].

6. Implications for Genetic Diversity in Grapevine: What We Are Not Capturing

If the use of a single reference genome leaves part of the genetic variation outside the analysis, the question is not only what is missing, but also how much of that absence may already be influencing the conclusions that are later taken for granted. This reflects a broader problem of representation in crop genomics, where part of the existing variation remains effectively unobserved because it is not adequately captured within the analytical model itself [16,21].

One immediate implication is that part of the gene repertoire may simply not appear in standard analyses. Within reference-based approaches, genes absent from the reference sequence become, in practice, invisible, or are treated as if they did not exist. This limitation is consistent with empirical observations of intra-varietal and clonal variability in grapevine [5]. Similar situations have been reported repeatedly in pangenomic studies across several crops, where gene presence/absence variation represents a substantial component of genetic diversity [12,16,21]. Recent grapevine data support the same conclusion. The Grapepan resource described by [10] revealed extensive gene-content variation across cultivated genomes, indicating that a substantial fraction of diversity remains outside what can be represented by a single reference sequence. This affects not only how diversity is quantified, but also how it is ultimately defined, since that definition depends on what can actually be compared within the selected framework.

A second consequence concerns the identification of functional variation. Part of the structural variation and differences in gene content may not be reflected in polymorphisms that are easily detected through conventional reference-based analyses, which limits the identification of variants with potential functional relevance [17,20,25]. As a consequence, complex traits such as stress response, local adaptation or specific quality attributes may end up being interpreted through only a partial view of the underlying variation. In grapevine, this could affect traits linked to aroma composition, including aroma-related compounds such as terpenes and norisoprenoids, responses to drought or heat stress, or genes associated with resistance to diseases such as downy or powdery mildew, whose genetic basis may not always be fully represented in reference genomes [10,14,24]. Evidence supporting this possibility is already beginning to emerge. Using a *Vitis* super-pangenome, ref. [14] identified structural variants associated with resistance to *Plasmopara viticola* and

highlighted candidate genes that would have been difficult to recover through conventional reference-based analyses alone. This may be particularly relevant in long-isolated viticultural systems, where continuous clonal propagation and local adaptation have favoured the accumulation of distinctive molecular profiles and intra-varietal divergence, as described in volcanic Atlantic vineyards such as Lanzarote [37].

There is also a less obvious effect. When diversity is interpreted through a reference genome, it is often assumed, implicitly, that most differences between genotypes are mainly explained by point variants. Yet, in grapevine, as in other crops, an important fraction of the variation may instead involve structural rearrangements or changes in gene content [7–9,16,21]. This type of variation remains only partially represented when analyses continue to rely predominantly on linear models.

Under these conditions, the consequence is not simply an underestimation of diversity, but a distorted representation of it. Certain forms of variation become overrepresented because they are easier to detect, whereas others remain comparatively underexplored. This directly affects the ability to establish reliable relationships between genotype and phenotype, which is central both to the interpretation of genetic diversity and to its practical use [16,21].

From an applied perspective, these limitations are not trivial. If part of the functional variation remains undetected, its incorporation into breeding programmes becomes substantially more difficult, for example, in the identification of genes associated with downy mildew resistance or in the modulation of aromatic profiles. In species such as grapevine, characterised by high intra-specific diversity and by selection under highly variable environmental conditions, this limitation may become especially relevant [10,14,16,21].

Taken together, these observations suggest that the main consequence of reference-based analyses may not be the simple omission of a fraction of genetic diversity. The problem is potentially more subtle. When some forms of variation are systematically easier to detect than others, the resulting picture of diversity may become uneven, emphasising certain genetic signals while underrepresenting others. Under these conditions, the challenge is not only to recover additional variants, but to reassess whether current interpretations of adaptation, phenotypic variation or genetic relationships are being built on a sufficiently complete representation of the available diversity. In this sense, pangenomics does not merely expand the catalogue of observable variation. It changes the context within which that variation is interpreted.

Overall, the issue goes beyond methodology alone. It also concerns the way genetic diversity is currently being observed and interpreted. As long as analyses continue to depend predominantly on a single reference genome, part of the available variation will remain difficult to identify, compare and ultimately use [16,21,22,35]. In this sense, the limitation is not only related to resolution itself, but to the perspective from which diversity is analysed. A single reference genome can organise variation efficiently, but it cannot fully represent the biological complexity that exists within grapevine populations.

7. Specific Challenges of Pangenomics in Grapevine

The development of pangenomics in grapevine depends less on generating additional genome assemblies than on how these data are integrated and made comparable, and on a set of biological and structural features that condition how these data can be analysed and interpreted. These factors do not prevent the application of pangenomic approaches, but they help explain why their integration has been slower than in other crops [18,22,38].

One of the most relevant elements is clonal propagation. Unlike species in which sexual reproduction dominates breeding programmes, in grapevine vegetative multiplication favours the progressive accumulation of somatic mutations. This process generates

intra-varietal diversity that is difficult to capture with conventional genomic schemes and complicates the distinction between stable structural variants and changes specific to particular clones [6,7,26]. In practice, this means that the identification of shared variants between genotypes may be partially obscured, or, in some cases, effectively masked, by recent somatic variation, making it harder to define consistent patterns at the population level.

The biological relevance of this process is illustrated by well-known somatic variants identified in traditional grapevine cultivars [39]. A classic example is Tempranillo Blanco, which originated through a spontaneous somatic mutation of Tempranillo Tinto and differs in several phenotypic and agronomic traits despite sharing most of its genetic background [40]. Similar cases have been reported in other cultivars and highlight how vegetative propagation can preserve and accumulate biologically relevant variation over long periods of time [7].

To this is added the high heterozygosity characteristic of many *Vitis vinifera* L. genotypes. This condition not only complicates the assembly of high-quality genomes, but also introduces a level of ambiguity that is not easily resolved in comparisons between individuals, particularly in structurally variable regions or in regions affected by gene presence/absence [8,9,24,36]. In this context, building robust pangenomes requires more demanding analytical strategies than in species with more homogeneous genomes. Operationally, this translates into greater uncertainty when defining the core genome and the variable fraction, especially when the number of analysed genotypes remains limited.

Another limiting factor is the scale of available genomic resources. Although the number of assemblies in grapevine has increased in recent years, it is still comparatively small, and unevenly distributed, compared to crops in which pangenomics is more established. This also affects how representative the included diversity is, since many local varieties, traditional materials and wild *Vitis* species remain poorly characterised at the genomic level [14,15].

The historical structure of diversity in grapevine introduces an additional layer of complexity. Domestication, clonal selection and historical exchanges of plant material have generated a network of genetic relationships that is difficult to represent using linear models, including cases where varietal identity and lineage remain only partially resolved, as discussed for the Malvasia group [41]. In this context, genetic variation does not only reflect recent selection processes, but also deeper evolutionary dynamics that may influence how variants are distributed across the genome [1,4,26].

Finally, there is a limitation related to data integration. Unlike in other crops, where pangenomes have progressively been incorporated into functional analysis platforms and breeding programmes, in grapevine these approaches are not yet part of standardised workflows. This makes genomic information harder to reuse, and even harder to compare across studies, and limits its connection with phenotypic and agronomic data, which is a necessary step if pangenomics is to have a real practical application [14,23]. As a consequence, the transfer of genomic results to applied contexts, such as varietal selection or breeding, remains limited, not so much because of a lack of identified variation, but because of the difficulty of integrating it into comparable and usable frameworks.

Taken together, these factors should not be interpreted only as technical obstacles, because they are not merely technical, but as elements that define the context in which pangenomics in grapevine must develop. Understanding these limitations is essential not only for interpreting current results correctly, but also for defining realistic strategies that enable progress towards a more complete and operational integration of genetic diversity in this species [18,22].

8. Future Perspectives: Towards an Operational Pangenomics in Grapevine

The development of operational pangenomics in grapevine will depend on how newly generated genomic resources are integrated into comparable and functionally interpretable frameworks. In this sense, the progress of the field appears to be shaped less by the volume of data than by how those data are structured, and, crucially, whether they can actually be compared [22,27].

A first line of development concerns expanding both the number and the diversity of genotypes included in grapevine pangenomes. In other crops, the incorporation of cultivated materials, traditional varieties and wild relatives has made it possible to capture fractions of variation that were previously not represented and to improve the characterisation of genetic diversity at a global scale [29,30]. In grapevine, this is particularly relevant given the coexistence of cultivated materials, local resources and wild species of the genus *Vitis*, which are still insufficiently represented in current genomic resources [15].

A second key direction is the development of more integrative pangenomic representations. Graph-based approaches have allowed, in different systems, for a more effective handling of structural variation and a better integration of this variation into comparative analyses, especially in complex genomes [31,33]. In species with high heterozygosity and structural diversity, such as grapevine, these approaches are likely to prove decisive for improving both analytical resolution and reproducibility.

Beyond methodological aspects, the most relevant advance is likely to be the integration of genomic variation with phenotype. In other crops, combining pangenomics with multi-omics approaches has made it possible to establish more direct links between genetic variants and complex traits, including productive and adaptive characteristics [25,34]. In grapevine, this type of integration remains at an early stage, arguably the main limitation at present, which limits the capacity to translate genetic diversity into functional and applicable knowledge [23].

From an applied perspective, incorporating pangenomics into breeding programmes represents one of the main opportunities, perhaps the most immediate one, in the field. The identification of structural variants, genes absent from reference genomes and genotype-specific regions can substantially expand the pool of variants available for selection [14,35]. In this context, the development of comparable and accessible resources will be a prerequisite for practical implementation.

At the same time, coordination between research groups and the standardisation of methodologies emerge as key factors. In other crops, the development of large-scale pangenomes has often been associated with collaborative initiatives that have enabled the generation of broad and comparable datasets, facilitating their reuse across different analytical contexts [18,36]. In grapevine, where data remain more fragmented, similar approaches could significantly accelerate the consolidation of the field.

One particularly important challenge for grapevine concerns the incorporation of local genetic resources, clonal diversity and underrepresented germplasm into future pangenomic frameworks. Unlike many annual crops, a substantial fraction of grapevine diversity has been maintained through centuries of vegetative propagation and is often represented by locally distributed cultivars, biotypes and clonal lineages that remain poorly characterised at the genomic level. This diversity is not only relevant from a conservation perspective. It may also contain variants associated with environmental adaptation, stress tolerance or quality-related traits that are difficult to detect within more restricted genomic representations. Future pangenomic resources will therefore need to capture diversity across multiple biological scales, from wild relatives and cultivated varieties to intra-varietal and regional variation, if they are to provide a realistic representation of the

genetic resources currently available for grapevine research and breeding. This point is especially relevant for grapevine because some of the most valuable diversity is not necessarily found in large, modern breeding populations. It may be present in old local cultivars, pre-phylloxera materials, long-isolated vineyard systems or clonal lineages that have been maintained vegetatively for decades, sometimes centuries. These materials are often difficult to place within conventional genomic frameworks, but they may contain variation linked to drought response, heat tolerance, phenology, disease resistance or berry composition. In a perennial crop facing increasingly unstable climatic conditions, this is not a secondary issue. Future pangenomic resources should therefore be designed not only to represent broad taxonomic diversity, but also the historical and agronomic diversity that has shaped grapevine cultivation.

An equally important question concerns what kind of diversity future pangenomic frameworks are expected to represent. Most efforts have focused on increasing the number of available genomes, which is clearly necessary. Yet, the examples discussed throughout this review suggest that representation may be as important as scale itself. A pangenome built from a large number of closely related genotypes does not necessarily capture the same biological diversity as one that incorporates wild relatives, local cultivars, long-isolated populations or intra-varietal variation. In grapevine, where historical diversification has occurred across highly heterogeneous environmental and cultural contexts, this distinction may prove particularly important. Future progress will therefore depend not only on expanding genomic resources, but also on ensuring that the diversity incorporated into those resources adequately reflects the biological complexity of the species.

Overall, the future of pangenomics in grapevine does not depend solely on describing genetic diversity with greater precision. The more fundamental challenge is to ensure that the diversity being represented corresponds, as closely as possible, to the biological diversity that actually exists within the species. From this perspective, pangenomics is not simply a strategy for generating larger genomic datasets. It is an attempt to reduce the gap between the diversity that exists and the diversity that can be observed, interpreted and ultimately used. Whether this objective can be achieved will depend less on the number of genomes available than on the capacity to connect genomic variation with phenotype, adaptation and practical applications [22]. The successive steps required to translate grapevine pangenomic resources into functional, conservation and breeding applications are summarised in Figure 3.

In grapevine, particular attention should be given to clonal diversity and long-term vegetative lineages, which represent a unique source of variation not typically addressed in annual crops.

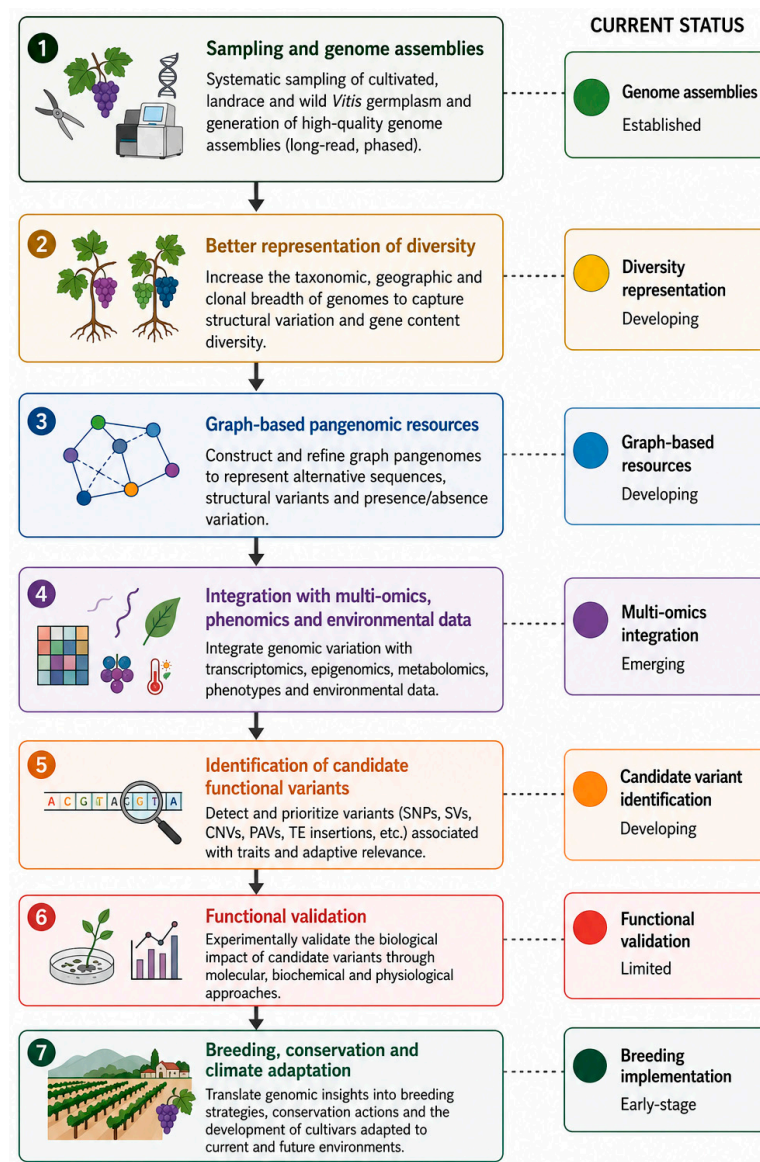


Figure 3. Conceptual roadmap towards operational pangenomics in grapevine. The workflow illustrates the successive stages required to translate genomic resources into practical applications, from sampling and genome assembly to diversity representation, graph-based pangenomic resources, multi-omics integration, functional validation and breeding implementation. The panel on the right summarises the current level of development of each stage based on the literature discussed in this review. Conceptual illustration generated with the assistance of AI-based image tools and subsequently edited by the authors.

9. Conclusions

The recent development of grapevine pangenomes and super-pangenomes has begun to change the way genetic diversity is represented and analysed in *Vitis vinifera* L. The studies reviewed here consistently show that structural variation and gene presence/absence variation constitute relevant components of grapevine diversity and that part of this variation is only partially captured when analyses rely on a single reference genome.

Current pangenomic resources have already provided a broader view of cultivated and wild *Vitis* diversity, revealing genomic regions, genes and structural variants that were previously difficult to detect. At the same time, the available evidence indicates that grapevine pangenomics is still in a transitional stage. Although important resources such as Grapepan and recent *Vitis* super-pangenomes have considerably expanded the

representation of diversity, the connection between genomic variation, phenotype and breeding applications remains limited when compared with several major crop species.

The literature also suggests that the future value of grapevine pangenomics will depend less on the continuous generation of additional genome assemblies than on the ability to integrate genomic resources with functional validation, transcriptomic information, phenotypic datasets and breeding programmes. In this sense, the main challenge is no longer demonstrating that additional diversity exists, but understanding how that diversity contributes to adaptation, stress tolerance, disease resistance and grape quality.

Particular attention should be given to the incorporation of underrepresented local germplasm, traditional cultivars and wild *Vitis* resources since these materials may contain genetic variation of increasing relevance under climate change scenarios. Expanding their representation within pangenomic frameworks could improve both the identification of adaptive alleles and the long-term conservation of grapevine genetic resources.

Ultimately, pangenomics should not be viewed simply as a more comprehensive catalogue of genomic variation. Its practical significance will depend on whether it can help translate previously hidden genetic diversity into improved biological understanding and, eventually, into tools that support grapevine conservation, climate adaptation and breeding.

Author Contributions: Conceptualisation, F.F.; methodology, F.F.; formal analysis, F.F.; investigation, F.F., L.D. and Q.L.-Y.; resources, F.F. and F.Z.; data curation, F.F. and Q.L.-Y.; writing—original draft preparation, F.F.; writing—review and editing, F.F., L.D., J.M.C. and F.Z.; visualisation, F.F. and Q.L.-Y.; supervision, F.Z. and J.M.C.; project administration, F.F. 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.

References

1. This, P.; Lacombe, T.; Thomas, M.R. Historical origins and genetic diversity of wine grapes. *Trends Genet.* **2006**, *22*, 511–519. [[CrossRef](#)] [[PubMed](#)]
2. 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](#)]
3. 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](#)]
4. 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](#)]
5. Fort, F.; Suárez-Abreu, L.R.; Lin-Yang, Q.; Deis, L.; Canals, J.M.; Zamora, F. Preliminary clonal characterization of *Malvasia volcanica* and *Listan prieto* by simple sequence repeat (SSR) markers in free-phylloxera volcanic vineyards (Lanzarote and Fuerteventura, Canary Islands, Spain). *Horticulturae* **2025**, *11*, 823. [[CrossRef](#)]
6. Allaby, R.G. Clonal crops and the role of structural variation in domestication. *Nat. Plants* **2019**, *5*, 915–916. [[CrossRef](#)] [[PubMed](#)]
7. Vondras, A.M.; Minio, A.; Blanco-Ulate, B.; Figueroa-Balderas, R.; Penn, M.A.; Zhou, Y.; Seymour, D.; Ye, Z.; Liang, D.; Espinoza, L.K.; et al. The genomic diversification of grapevine clones. *BMC Genom.* **2019**, *20*, 972. [[CrossRef](#)] [[PubMed](#)]
8. Liang, Z.; Duan, S.; Sheng, J.; Zhu, S.; Ni, X.; Shao, J.; Liu, C.; Nick, P.; Du, F.; Fan, P.; et al. Whole-genome resequencing of 472 *Vitis* accessions for grapevine diversity and demographic history analyses. *Nat. Commun.* **2019**, *10*, 1190. [[CrossRef](#)] [[PubMed](#)]
9. 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](#)]
10. 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](#)]

11. Tettelin, H.; Masignani, V.; Cieslewicz, M.J.; Donati, C.; Medini, D.; Ward, N.L.; Angiuoli, S.V.; Crabtree, J.; Jones, A.L.; Durkin, A.S.; et al. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial pan-genome. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13950–13955. [[CrossRef](#)] [[PubMed](#)]
12. Golicz, A.A.; Batley, J.; Edwards, D. Towards plant pangenomics. *Plant Biotechnol. J.* **2016**, *14*, 1099–1105. [[CrossRef](#)] [[PubMed](#)]
13. Cochetel, N.; Minio, A.; Guercio, 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](#)]
14. Guo, L.; Wang, X.; Ayhan, D.H.; Rhaman, M.S.; Yan, M.; Jiang, J.; Wang, D.; Zheng, W.; Mei, J.; Ji, W.; et al. Super pangenome of *Vitis* empowers identification of downy mildew resistance genes for grapevine improvement. *Nat. Genet.* **2025**, *57*, 741–753. [[CrossRef](#)] [[PubMed](#)]
15. Cantu, D.; Massonnet, M.; Cochetel, N. The wild side of grape genomics. *Trends Genet.* **2024**, *40*, 601–612. [[CrossRef](#)] [[PubMed](#)]
16. Schreiber, M.; Jayakodi, M.; Stein, N.; Mascher, M. Plant pangenomes for crop improvement, biodiversity and evolution. *Nat. Rev. Genet.* **2024**, *25*, 563–577. [[CrossRef](#)] [[PubMed](#)]
17. Gao, L.; Gonda, I.; Sun, H.; Ma, Q.; Bao, K.; Tieman, D.M.; Burzynski-Chang, E.A.; Fish, T.L.; Stromberg, K.A.; Sacks, G.L.; et al. The tomato pangenome uncovers new genes and rare alleles associated with fruit flavor. *Nat. Genet.* **2019**, *51*, 1044–1051. [[CrossRef](#)] [[PubMed](#)]
18. He, W.; Li, X.; Qian, Q.; Shang, L. The developments and prospects of plant super-pangenomes. *Plant Commun.* **2025**, *6*, 101230. [[CrossRef](#)] [[PubMed](#)]
19. Montenegro, J.D.; Golicz, A.A.; Bayer, P.E.; Hurgobin, B.; Lee, H.; Chan, C.-K.K.; Visendi, P.; Lai, K.; Doležel, J.; Batley, J.; et al. The pangenome of hexaploid bread wheat. *Plant J.* **2017**, *90*, 1007–1013. [[CrossRef](#)] [[PubMed](#)]
20. Zhao, Q.; Feng, Q.; Lu, H.; Li, Y.; Wang, A.; Tian, Q.; Zhan, Q.; Lu, Y.; Zhang, L.; Huang, T.; et al. Pan-genome analysis highlights the extent of genomic variation in rice. *Nat. Genet.* **2018**, *50*, 278–284. [[CrossRef](#)] [[PubMed](#)]
21. Bayer, P.E.; Golicz, A.A.; Scheben, A.; Batley, J.; Edwards, D. Plant pan-genomes are the new reference. *Nat. Plants* **2020**, *6*, 914–920. [[CrossRef](#)] [[PubMed](#)]
22. Loegler, V.; Friedrich, A.; Schacherer, J. Dynamics of genome evolution in the era of pangenome analysis. *Cell Genom.* **2026**, *6*, 101067. [[CrossRef](#)] [[PubMed](#)]
23. Liu, Y.; Dong, Y.; Fang, X.; Wang, X.; Li, J.; Zhang, T.; Long, Q.; Su, Y.; Shi, X.; Su, X.; et al. Pangenomics and single-cell transcriptomics uncover the genetic basis of continuous bearing trait in grapevine. *Hortic. Res.* **2025**, *12*, uhaf228. [[CrossRef](#)] [[PubMed](#)]
24. Ma, Z.; Xu, X.; Peng, W.; Zhang, T.; Chen, Z.; Cao, S.; Zhang, F.; Wang, Y.; Xiao, H.; Zhang, Y.; et al. Population genomics reveals association of transposable element variants with climatic adaptation in wild Amur grape. *Nat. Commun.* **2026**, *17*, 3213. [[CrossRef](#)] [[PubMed](#)]
25. Morris, G.P.; Harder, A.M.; Healey, A.L.; McLaughlin, C.M.; Rifkin, J.L.; Cruet-Burgos, C.; Jenkins, J.W.; Shu, S.; Spiekerman, J.J.; VanGessel, C.J.; et al. A sorghum pangenome reference improves global crop trait discovery. *Nature* **2026**, *652*, 1245–1253. [[CrossRef](#)]
26. 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](#)]
27. Ruperao, P.; Rangan, P.; Shah, T.; Sharma, V.; Rathore, A.; Mayes, S.; Pandey, M.K. Developing pangenomes for large and complex plant genomes. *J. Adv. Res.* **2025**, *78*, 29–46. [[CrossRef](#)] [[PubMed](#)]
28. Liu, Y.; Tian, Z. Super graph-based pan-genome: Bringing rice functional genomics into a new dawn. *Mol. Plant* **2022**, *15*, 1409–1411. [[CrossRef](#)] [[PubMed](#)]
29. Shang, L.; Li, X.; He, H.; Yuan, Q.; Song, Y.; Wei, Z.; Lin, H.; Hu, M.; Zhao, F.; Zhang, C.; et al. A super pan-genomic landscape of rice. *Cell Res.* **2022**, *32*, 878–896. [[CrossRef](#)] [[PubMed](#)]
30. Thudi, M.; Mascher, M.; Jayakodi, M. Pangenome charts the genomic path for wheat improvement. *Trends Plant Sci.* **2025**, *30*, 687–689. [[CrossRef](#)] [[PubMed](#)]
31. Zhou, Y.; Zhang, Z.; Bao, Z.; Li, H.; Lyu, Y.; Zan, Y.; Wu, Y.; Cheng, L.; Fang, Y.; Wu, K.; et al. Graph pangenome captures missing heritability and empowers tomato breeding. *Nature* **2022**, *606*, 527–534. [[CrossRef](#)] [[PubMed](#)]
32. Zhao, X.; Yu, J.; Zhang, J.; Sun, H.; Wu, S.; Zhao, J.; Zhou, Y.; Hammar, S.A.; Lin, Y.-C.; Zhang, Z.; et al. Graph-based pangenome reveals structural variation dynamics during cucumber breeding. *Nat. Genet.* **2026**, *58*, 643–654. [[CrossRef](#)] [[PubMed](#)]
33. Du, Z.-Z.; He, J.-B.; Jiao, W.-B. Plant graph-based pangenomics: Techniques, applications, and challenges. *aBIOTECH* **2025**, *6*, 361–376. [[CrossRef](#)] [[PubMed](#)]
34. Yuan, X.; Jiang, X.; Zhang, M.; Wang, L.; Jiao, W.; Chen, H.; Mao, J.; Ye, W.; Song, Q. Integrative omics analysis elucidates the genetic basis of seed weight and oil content in soybean. *Plant Cell* **2024**, *36*, 2160–2175. [[CrossRef](#)] [[PubMed](#)]
35. Raza, A.; Li, Y.; Jan, F.; Fernandez, C.G.T.; Mir, R.R.; Hu, Z.; Varshney, R.K. From the genome to super-pangenome: A new paradigm for accelerated crop improvement. *npj Sci. Plants* **2026**, *2*, 4. [[CrossRef](#)]

36. Jayakodi, M.; Lu, Q.; Pidon, H.; Rabanus-Wallace, M.T.; Bayer, M.; Lux, T.; Guo, Y.; Jaegle, B.; Badea, A.; Bekele, W.; et al. Structural variation in the pangenome of wild and domesticated barley. *Nature* **2024**, *636*, 654–662. [[CrossRef](#)] [[PubMed](#)]
37. Fort, F.; Marsal, G.; Mateo-Sanz, J.M.; Pena, V.; Canals, J.M.; Zamora, F. Molecular characterisation of the current cultivars of *Vitis vinifera* L. in Lanzarote (Canary Islands, Spain) reveals nine individuals which correspond to eight new varieties and two new sports. *Oeno One* **2022**, *56*, 281–295. [[CrossRef](#)]
38. Wang, Y.; Ding, K.; Li, H.; Kuang, Y.; Liang, Z. Biography of *Vitis* genomics: Recent advances and prospective. *Hortic. Res.* **2024**, *11*, uhae128. [[CrossRef](#)] [[PubMed](#)]
39. Fernández, L.; Torregrosa, L.; Segura, V.; Bouquet, A.; Martínez-Zapater, J.M. Transposon-induced gene activation as a mechanism generating cluster shape somatic variation in grapevine. *Plant J.* **2010**, *61*, 545–557. [[CrossRef](#)] [[PubMed](#)]
40. Carbonell-Bejerano, P.; Royo, C.; Torres-Pérez, R.; Grimplet, J.; Fernandez, L.; Franco-Zorrilla, J.M.; Lijavetzky, D.; Baroja, E.; Martínez, J.; García-Escudero, E.; et al. Catastrophic unbalanced genome rearrangements cause somatic loss of berry color in grapevine. *Plant Physiol.* **2017**, *175*, 786–801. [[CrossRef](#)] [[PubMed](#)]
41. Fort, F.; Suárez-Abreu, L.R.; Lin-Yang, Q.; Deis, L.; Canals, J.M.; Zamora, F. Origin and possible members of the ‘Malvasia’ family. *Horticulturae* **2025**, *11*, 561. [[CrossRef](#)]

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.