HAL open science

Distribution of prophages in the Oenococcus oeni species

Olivier Claisse, Ame Chaïb, Fety Jaomanjaka, Cécile Philippe, Yasma Barchi, Patrick Lucas, Claire Le Marrec

To cite this version:

Olivier Claisse, Ame Chaïb, Fety Jaomanjaka, Cécile Philippe, Yasma Barchi, et al.. Distribution of prophages in the Oenococcus oeni species. Microorganisms, 2021, 9 (4), pp.856. 10.3390/microorganisms . hal-03307703

HAL Id: hal-03307703

https://hal.inrae.fr/hal-03307703

Submitted on 29 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

microorganisms

MDPI

Article

Distribution of Prophages in the *Oenococcus oeni* Species

Olivier Claisse, Amel Chaïb, Fety Jaomanjaka, Cécile Philippe, Yasma Barchi, Patrick M. Lucas and Claire Le Marrec *

Unité de Recherche Œnologie, Bordeaux INP, University of Bordeaux, INRAE, ISVV, F-33882 Bordeaux, France; olivier.claisse@u-bordeaux.fr (O.C.); bounoise68@hotmail.fr (A.C.); jaomazava@yahoo.fr (F.J.); cecilemphilippe@gmail.com (C.P.); barchiyasma34@gmail.com (Y.B.); patrick.lucas@u-bordeaux.fr (P.M.L.)
* Correspondence: clehenaff@enscbp.fr; Tel.: +33-557-575-831

Abstract: *Oenococcus oeni* is the most exploited lactic acid bacterium in the wine industry and drives the malolactic fermentation of wines. Although prophage-like sequences have been identified in the species, many are not characterized, and a global view of their integration and distribution amongst strains is currently lacking. In this work, we analyzed the complete genomes of 231 strains for the occurrence of prophages, and analyzed their size and positions of insertion. Our data show the limited variation in the number of prophages in *O. oeni* genomes, and that six sites of insertion within the bacterial genome are being used for site-specific recombination. Prophage diversity patterns varied significantly for different host lineages, and environmental niches. Overall, the findings highlight the pervasive presence of prophages in the *O. oeni* species, their role as a major source of within-species bacterial diversity and drivers of horizontal gene transfer. Our data also have implications for enhanced understanding of the prophage recombination events which occurred during evolution of *O. oeni*, as well as the potential of prophages in influencing the fitness of these bacteria in their distinct niches.

Keywords: bacteriophage; *Oenococcus oeni*; malolactic fermentation; integrase; recombination

check for updates

Citation: Claisse, O.; Chaïb, A.; Jaomanjaka, F.; Philippe, C.; Barchi, Y.; Lucas, P.M.; Le Marrec, C. Distribution of Prophages in the *Oenococcus oeni* Species. *Microorganisms* **2021**, *9*, 856. https://doi.org/10.3390/microorganisms9040856

Academic Editor: Hanna M. Oksanen

Received: 23 March 2021
Accepted: 14 April 2021
Published: 16 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 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 (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The wine-making process starts with the selection of the fruit and the fermentation of sugars into alcohol by yeasts. In most red and dry white wines, malolactic fermentation (MLF), which reduces acidity, increases microbial stability and creates good-quality grape wine, is also a required step. MLF is largely driven by the lactic acid bacterium (LAB) *Oenococcus oeni* [1]. This bacterium is rarely detected in the natural environment, including the surface of grape berries [2], and its population only increases after crushing. At this step, *O. oeni* is part of a complex microbiota that still comprises different LAB. Wine conditions become progressively harsher for most bacteria, except for *O. oeni*, which becomes the sole detectable LAB species isolated from wine when MLF typically occurs, and is indeed the best adapted species to the combined inhibitory effects of low pH, low oxygen, high alcohol, polyphenolic compounds and sulfur dioxide contents [2]. However, intra-species diversity is reported in the *O. oeni* species and large phenotypic differences among indigenous strains corresponding to variable capacities to withstand stringent wine compositions and impact aromas have been repeatedly observed. This has become the focus of more detailed studies, with the aim to explain microbial phenotypes on a genomic scale. To this aim, collections of strains have been assembled worldwide, from all wine regions, and also from other niches supporting the growth of *O. oeni* such as apple cider and kombucha, a fermented tea containing very little alcohol [2]. As from 2005 [3], an enormous amount of information has been progressively obtained from whole genome sequencing, showing that *O. oeni* is a highly diverse species, with genetic adaptation to fermented beverages [4–7]. Such a specialization to life in wine is visible in the small (1.6 to 2.2 Mbp) genome of *O. oeni* because the species has lost many metabolic abilities due to its adaptation to an environment rich

in amino acids, vitamins and other nutrients [7]. The population structure revealed by MLST and phylogenomic analyses currently recognizes four main phylogenetic groups (referred to as groups A, B, C and D) [7–9]. Group A exclusively contains wine strains, and is divided into subgroups which are associated with certain types of wine, such as white wines of Burgundy or Champagne, supporting the idea that group A strains are the most domesticated to wine [10]. Groups B and C contain both cider and wine strains [7]. More recently, strains originating from kombucha tea were shown to form a fourth group named phylogroup D, which may have preceded the advent of other groups [9].

A key factor in the rapid evolution and adaptation of the strains is undoubtedly the hypermutability of the *O. oeni* species [11]. As reported for a few other species, the absence of the ubiquitous mismatch repair (MMR) genes *mutS* and *mutL* results in the accumulation of spontaneous errors in DNA replication, or in reduced stringency in recombination, thus generating high levels of polymorphism, increasing adaptation to environmental fluctuation [12]. The mutator phenotype also promotes horizontal distribution of essential genes for survival, and therefore increases the genetic innovation rate, eliciting environmental adaptation [12]. The species is also devoid of a CRISPR-Cas system, the bacterial immunity defense mechanism against foreign DNA [2–11]. Given these peculiar characteristics, it is therefore not surprising that mobile genetic elements such as insertion sequences (ISs) [13], plasmids [14] and DNA of (pro)phage origin [15] have been shown to constitute a significant portion of the *O. oeni* pangenome. The observed prevalence of prophages in bacterial genomes is indeed in agreement with previous cultivation-based assessments of lysogeny reported in the species [16].

Additional insights into the diversity of O. oeni prophages was obtained through the analysis of two major constituents of the site-specific recombination (SSR) unit, corresponding to the phage integrases and attachment sites [4,17]. Alignment of the integrase protein sequences was shown to cluster prophages into four major groups (Int_A to Int_D), which were related to the integration site in the host chromosome [17]. Thus far, all groups, except Int_C, contain mitomycin C-inducible prophages. Accordingly, no Int_C phage has been isolated as free phage particles during field campaigns targeting samples from grapes, wines and materials [17,18].

The abundant literature on diverse ecosystems fosters further effort to generate deeper knowledge of lysogeny in the *O. oeni* species. In particular, there is a convergence of case studies towards a role of prophages in providing their host bacterium with ecologically significant traits [19,20]. The later include immunity against specific phage attack [21], virulence [22] or specific adhesion properties through the modification of the cell surface matrix [20,23], thereby improving bacterial survival under changing conditions. Ecosystem-level consequences of lysogeny have also been documented. For example, prophages modulate the interactions between microbial symbionts and their insect or human hosts [24,25]. The current trend toward understanding phage–host interactions on a global scale also makes a valuable contribution on how phages influence bacterial diversification and ecotype formation [26,27]. The deciphering of the mechanisms of ecological divergence in the *Oenococcus* genus is highly topical, as three novel species have recently been proposed during exploration of different fermented beverages [2]. This has expanded our knowledge of the horizons of their habitat to other niches such as alcohol production from sugar cane for *O. alcoholitolerans* [28], shochu distillates for *O. kitaharae* [29], ciders for *O. oeni* and *O. sicerae* [30] and kefirs for *O. sicerae* [31]. Thanks to the recent increase in the number and diversity of sequenced genomes, we investigated the distribution, organization and insertion of prophages in *O. oeni* and its related species. Our data will help in understanding the evolutionary trajectories of strains and their phages in fermented beverages, and can be further used to explore whether prophages play a role during the adaptation of their host to wine-making conditions.

2. Material and Methods

2.1. Data Collection and Identification of Candidate Prophage Sequences in Complete Genomes of O. oeni

A total of 231 publicly available genome sequences were obtained from the National Center for Biotechnology Information NCBI, https://www.ncbi.nlm.nih.gov (accessed on 2 January 2021) (Table S1). Corresponding strains have been collected worldwide from different fermented beverages (red, dry and sweet white wines, sparkling wines and more recently cider and kombucha tea). Candidate prophage-like elements were identified using the RAST pipeline which provides an automated approach to phage genome annotation [32]. As suggested by others [33], we performed a manual inspection of the sequences for the presence of signature sequences: attachment sites (*att*), gene(s) encoding integrase(s), terminases(s), transposases(s), genes coding for structural viral proteins and the sequences of prophage integration sites. In particular, a multi-FASTA file containing the four publicly available integrase protein sequences described among temperate oenophages was used as a reference to blast the draft genomes. In addition, the prophage search tool PHASTER was also used [34].

2.2. Comparisons and Phylogenetic Analyses

For specific ORF phylogenetic analyses, protein sequences were aligned using ClustalOmega at https://toolkit.tuebingen.mpg.de/ (accessed on 2 January 2021).

Phage genome comparisons were conducted using the Genome-BLAST Distance Phylogeny (GBDP) method using VICTOR under settings recommended for prokaryotic viruses [35] at: http://ggdc.dsmz.de/phylogeny-service.php (accessed on 2 January 2021). The resulting intergenomic distances were used to infer a balanced minimum evolution tree with branch support via FASTME including SPR postprocessing for each of the formulas D0, D4 and D6, respectively. Branch support was inferred from 100 pseudo-bootstrap replicates each. Trees were rooted at the midpoint and visualized with FigTree [36] as already described [37].

2.3. PCR

Bacterial strains CRBO11105 and CRBO14210 (accession numbers LKSR01 and LKRV01) were obtained from the Centre de Ressources Biologiques Oenologiques (CRB Oeno; ISVV, Villenave d'Ornon, France). Genomic DNAs from were extracted from cultures by using the Whatman® FTA Clone Saver card technology (Whatman, Sigma-Aldrich, France). PCR amplification reactions were achieved in a 25 µL final volume with 0.2 µM of each primer. The Taq 5X Master Mix kit (New England, Biolabs, Evry, France) was used according to the manufacturer's recommendations. A BioRad i-Cycler was used for amplification. Bacterial DNA was introduced in the reaction mixture as 1.2 mm FTA discs which were punched out of the center of FTA cards using a Uni-Core Punch (Qiagen Cat. No. WB100028, Whatman®, Fisher Scientific, Schwerte, Germany) and transferred to PCR tubes. The pair of primers targeting the $attB_B$ site was described earlier [17].

2.4. Induction with MC

For the induction of phages, mitomycin C was used as the inducing agent (1 µg/mL), with bacterial culture on modified MRS. Overnight cultures were diluted 10-fold in 10 mL of fresh broth, grown to an optical density at 600 nm (OD_{600}) of 0.2 to 0.3 prior to the addition of inducing agent and incubated for 24 h. OD_{600} was measured periodically.

3. Results and Discussion

3.1. The Genome of O. oeni Is Replete with Putative Prophages

The systematic interrogation of 231 complete genomes assessed from GenBank (NCBI) resulted in the discovery of seemingly intact prophages in 134 strains of *O. oeni* (58% of the bacterial genomes studied) (Table 1).

Table 1. Lysogeny in a set of 231 strains representing the four described phylogroups of *O. oeni* and different combinations of prophage carriage.

Distribution and Type of Prophages	Phylogroups in the *O. oeni* Species			
	A	**B**	**C**	**D**
Number of strains analyzed	174	31	21	5
Number of mono- and poly-lysogens	105	17	8	4
Lysogeny (%)	60	55	38	80
Mono-lysogens (n = 86) and prophage types				
Int_A	22	2	0	0
Int_B	14	1	1	0
Int_C	20	0	0	0
Int_D	3	12	4	0
Int_E	0	0	3	0
Int_F	0	0	0	4
Poly-lysogens (n = 48) and prophage combinations				
Int_{AB}	3	1	0	0
Int_{AC}	17	0	0	0
Int_{AD}	8	0	0	0
Int_{BC}	14	0	0	0
Int_{ABC}	1	1	0	0
Int_{ACD}	3	0	0	0
Phage remnants (PRs) in strains				
PR_C	174	0	9	0
PR_F	0	1	1	4
PR_G	0	0	0	4
PR_H	18	6	12	5
PR_I	9	14	15	0

Lysogens originated from all types of beverages (wines, ciders, kombucha), and belonged to all four lineages described in *O. oeni*, including the less numerously represented phylogroups associated with cider (B and C) and kombucha (D). Lysogens mostly harbored a single prophage (64.2%; n = 86). The 48 poly-lysogens contained two (n = 43) or three distinct prophages (n = 5) (Table 1).

This first survey yielded a total of 187 prophages. They showed a genome length ranging from 35 kb to 46.2 kb which, on average, and comprised between 2 and 6.7% of the host chromosome. All prophages encoded identifiable phage-specific functions such as integrases and terminases and tail-associated, portal-associated and lysis-associated proteins. The presence of a large number of predicted proteins characteristic of tailed phages (e.g., terminase, tape measure protein, tail formation and baseplate-related proteins) is consistent with previous observations by transmission electronic microscopy showing that oenophages display the Siphoviridae morphology [16–18]. Prophages also demonstrated well-conserved patterns in genome organization. Starting with the gene encoding the integrase, the following order was observed: lysogeny module followed by modules for replication, DNA packaging, head morphogenesis, tail, lysis and finally lysogenic conversion.

During a previous screening and analysis of prophages in a set of 42 publicly available genomes, we reported the existence of four distinct groups of temperate oenophages

(Int_{A-D}) [17]. The properties shared by all known members of a group included: (1) high identity of the integrase sequences (>98% at the amino acid level) and (2) tropism for a specific bacterial attachment site corresponding to the 3′ end of a tRNA gene [17]. To get a broader view, we extended our in silico analyses to the 187 newly retrieved prophages. The clustering of the majority of their integrase sequences into the four proposed groups (96%, $n = 180$) was confirmed. All identified Int_A, Int_C and Int_D integrases were observed to preferentially drive prophage integration into their expected cognate site, encompassing the 3′ end of the $tRNA^{Glu\text{-}0506}$ ($attB_A$), $tRNA^{Lys\text{-}0685}$ ($attB_C$) and $tRNA^{Leu\text{-}1359}$ ($attB_D$) sequences (gene numbers are those present in the PSU1 reference strain), respectively (Figure 1).

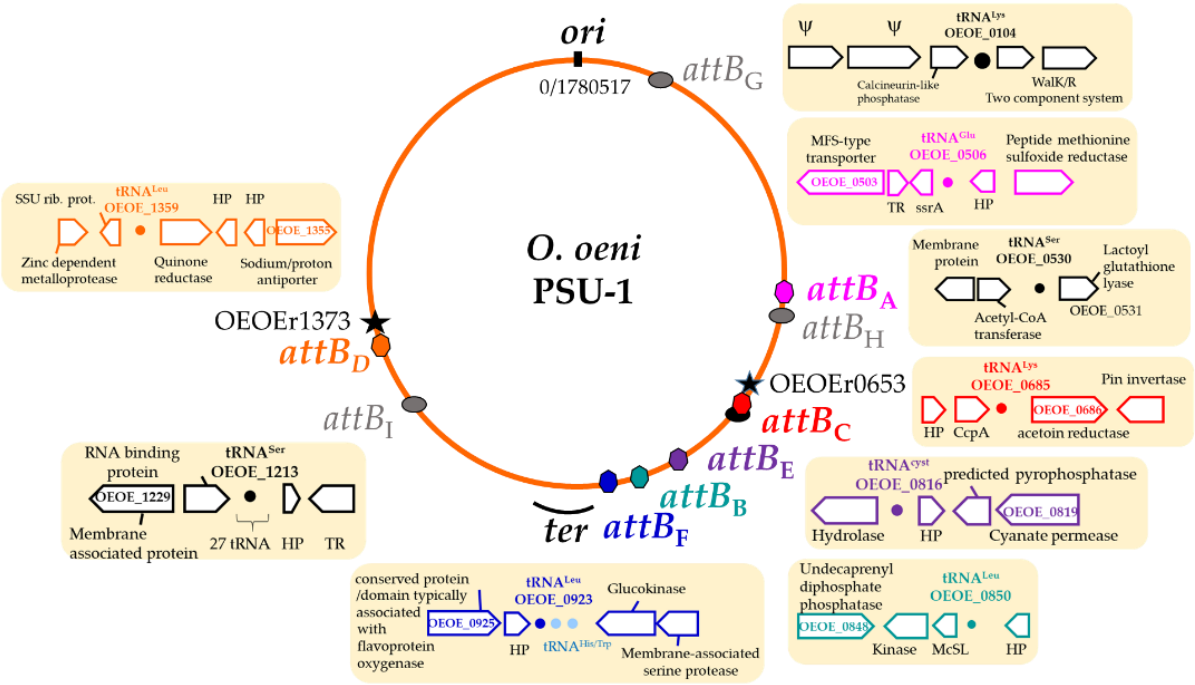

Figure 1. Site-specific recombination of temperate oenophages in *O. oeni.* The locations of the attachments sites used by prophages (color) and phage remnants (gray) are shown as a generalized chromosome backbone based upon the PSU-1 genome. Locations of the *oriC*, *Ter* site-containing region and rRNA loci (stars) are indicated. Genes flanking all *attB* sites are indicated.

Intriguing features were observed for some Int_B prophages. First, we confirmed that integration at the previously described $tRNA^{Leu\text{-}0851}$ ($attB_B$) site is widespread in the group, as this applied to 78% of the Int_B prophages [17]. In contrast, 22% showed an unusual localization outside the $attB_B$ site. Instead, recombination was shown to occur at the 3′ end of the $tRNA^{Leu\text{-}0923}$ gene, located 66 kb away (Figure 1). As this situation occurred in strains CRBO11105 and CRBO14210 from our strain collection, experimental verification of such an unusual integration site was done by PCR. We observed the failure to generate a PCR product using the bacterial chromosomal DNAs as a template and PCR primers flanking the $tRNA^{Leu\text{-}0923}$ gene, while an amplicon corresponding to the *attL* junction between bacterial and phage DNA was amplified in both strains. The Int_B phages were therefore integrated at the $tRNA^{Leu\text{-}0923}$ site in strains CRBO11105 and CRBO14210. Accordingly, the intact $attB_B$ region spanning the $tRNA^{Leu\text{-}0851}$ gene was amplified, confirming the absence of prophage at this site in both strains. It is therefore shown that a few Int_B prophages become integrated at a 63 bp secondary *att* site containing the 3′ end of the $tRNA^{Leu\text{-}0923}$ gene in *O. oeni*. This finding is in accordance with previous data suggesting that the Int_B phage 10MC could use the $tRNA^{Leu\text{-}0923}$ gene as a secondary site during lysogenization assays of the MCW strain [38].

Our genomic survey also indicated that seven of the 187 newly retrieved prophages harbored integrases with lower pairwise sequence identity (60–92%) with the existing Int_{A-D} proteins (Table 2).

Table 2. Nucleotide and amino acid identities in pairwise comparisons amongst the individual genes of phage integrases in *O. oeni*.

Integrase Types	Int_A	Int_B	Int_C	Int_D	Int_E
Int_B	69 [a]/62 [b]				
Int_C	68/60	84/86			
Int_D	74/66	68/61	68/59		
Int_E	66/61	72/69	72/68	66/60	
Int_F	70/61	88/92	88/92	69/60	73/68

Representative prophage integrase sequences for each group were obtained from the following *O. oeni* strains: Int_A from IOEB0608, Int_B from IOEBB10, Int_C from IOEBS28, Int_D from IOEB9805, Int_E from CRBO1384 and Int_F from UBOCC315001. [a] means: nucleotide identity. [b] means amico acid identity.

Three prophages harbored a 100% conserved integrase sequence, which shared 60 to 69% aa identity with the sequences affiliated to the four existing groups. This novel integrase, named Int_E, was associated with a 39.5 kb and a 38.9 kb prophage in the cider strains CRBO1384 and CRBO1389, respectively. It was also harbored by a slightly larger 46.3 kb prophage in strain AWRIB663 originating from wine. All three Int_E prophages targeted a 11 bp $attB_E$ core sequence to lysogenize their host, which consisted of the 3′ end of the unique *cysteinyl* $tRNA^{0816}$ gene in *O. oeni* [3]. The use of a $tRNA^{Cys}$ as a target for phage integration is rare amongst phages of LAB, but has been reported for a few phages such as TPW22 [39] and PH15 [40].

A second novel integrase, named Int_F, was associated with a 41.7 kb resident prophage, found in four out of five strains isolated from kombucha tea. Int_F shared a higher protein sequence identity (92%) with Int_B and Int_C integrases than with the three other groups (60–61%) (Table 2). Surprisingly, the site targeted by Int_F prophages ($attB_F$) was the same 63 bp *att* sequence containing the 3′ end of the $tRNA^{Leu\text{-}0923}$ gene, previously described as a putative alternative site for Int_B prophages (Figure 1). Hence, prophages with distinct integrases (Int_B or Int_F) can integrate at the same $tRNA^{Leu\text{-}0923}$ site in the chromosome of *O. oeni*.

Finally, we used mitomycin C (MC) to induce the prophages present in a set of representative strains from the CRB Oeno collection. We first tested 11 mono-lysogens (seven Int_A; two Int_B; two Int_D). We observed a clear lysis for nine strains, and the results obtained for strain IOEB0608 are shown in Figure S1. The obtained lysates formed plaques on strain IOEB277, suggesting that they contain functional phages (results not shown). The strains showing no lysis upon MC addition were IOEB0502 and C23, which harbored an Int_A and an Int_B phage, respectively. Lastly, we also tested the S28 double lysogen (Int_A-Int_C). The strain lysed in the presence of MC and plaques were also produced on the sensitive host. Phages were recovered from 20 individual plaques and characterized by PCR with primers targeting the int_A and int_C genes [17]. Only Int_A phages were detected, suggesting that Int_C phages do not form active phage particles or that Int_A is the faster phage which dictates lysis time.

3.2. Phage Remnants Also Use tRNA Sites in O. oeni

Degeneration of prophages was observed in *O. oeni* and eight distinct phage-related genomic islands were identified. Their grounding was further supported by the observations of reduced sizes (5.2 kb to 19.5 kb), defective integrase proteins (PR_C, $PR_{F1–2}$, PR_{I2}), the lack of other key phage genes (notably those involved in morphogenesis), the presence of transposase sequences (PR_{H2}, PR_{I1}) and degenerate *att* sites and/or larger sequence variability reflected by the presence of neighboring pseudogenes (PR_G, PR_F) (Figures 1 and 2). Such phage-related genomic islands were considered as prophage remnants (PRs).

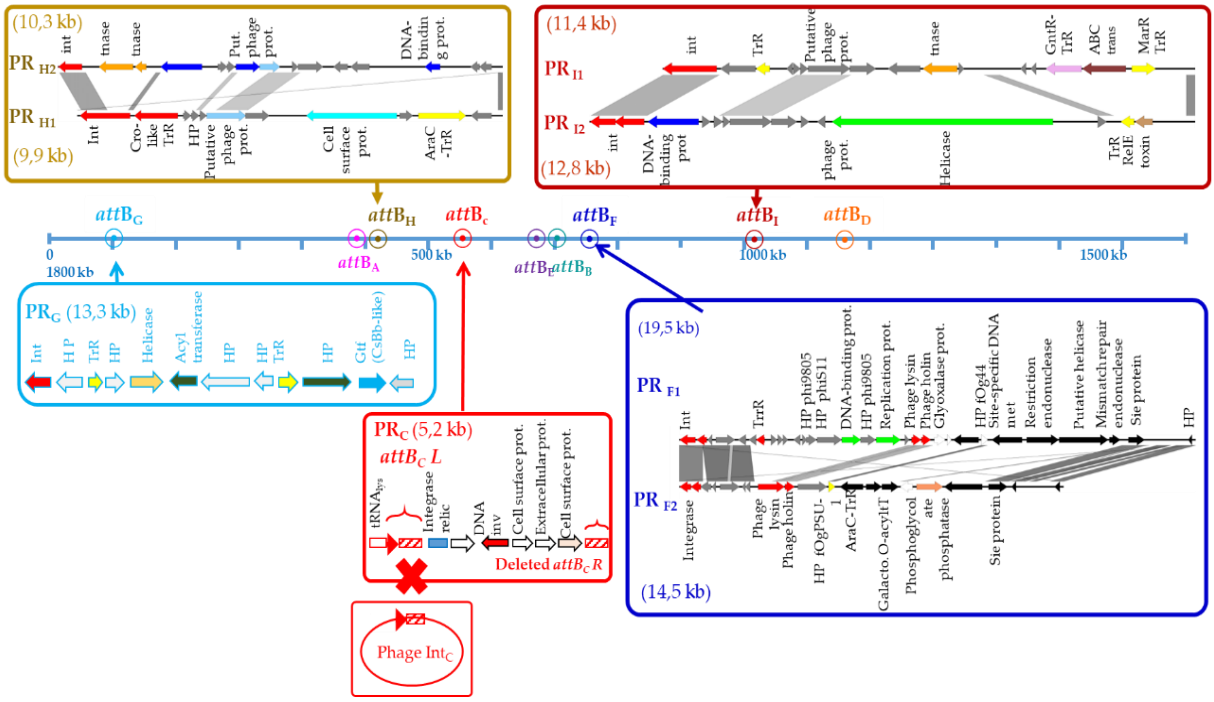

Figure 2. Localization and genetic organization of prophage remnants in the genome of *Oenococcus oeni*. Genes encoding hypothetical proteins are represented in gray.

PR_G was rare and found in 2.6% of the tested strains. The prevalence of PR_H and PR_I elements was similar (17.7% and 16.4%, respectively). Only PR_C was frequently observed in the species (79.2%) (Table 1).

PRs were consistently integrated in five precise locations within the host genome, which also corresponded to tRNA genes (Figures 1 and 2). Three PRs (PR_C and $PR_{F1/F2}$) were found at the previously characterized host sites used for the SSR of Int_C and Int_F phages, respectively [17]. However, the resident PRs and prophage sequences shared no homologies. The sequences flanking the five others (PR_G, $PR_{H1/2}$ and $PR_{I1/I2}$) differed from the sites of the known prophages of *O. oeni*, and were further named $attB_G$ (OEOE_t0104; $tRNA^{Lys}$), $attB_H$ (OEOE_t0530; $tRNA^{Ser}$) and $attB_I$ (OEOE_t1213; $tRNA^{Ser}$) (Figure 1). The length of the attachment sequences required for SSR were 61, 22 and 19 bp, respectively.

Annotation results indicated that a total of 16% of 122 ORFs found on all eight PRs were assigned to phage protein functions (Figure 2). Most were found in the larger PR_H regions (14.5 kb to 19.5 kb). Interestingly, only PR_{F1} and PR_{F2} still encode endolysin and holin genes. Several PR-related genes specified surface proteins of unknown function (PR_C, PR_{H1}), as well as putative phage resistance mechanisms (Sie, restriction–modification systems) (PR_{F1}, PR_{F2}, PR_G, PR_I). This suggests that PRs are not inactive DNA remnants and may encode traits useful for the host, as suggested for prophages [17]. In addition, the presence of phage resistance mechanisms in four PRs is an indicator of phage traffic and induced pressure during the making of fermented beverages.

3.3. Chromosomal Location and Genomic Context of Prophages and PR

As observed in many bacterial models, prophage-related regions (complete and PRs) are not randomly distributed in the *O. oeni* genome. Their presence was rare on the origin-proximal half of the *O. oeni* chromosome. Only PR_G was integrated near the origin of replication, where genes could be, on average, more expressed [41]. There was a significant enrichment in the absolute number of phages and PRs on the right replichore, where phage abundance increased along the *ori–ter* axis (Figure 1). Interestingly, four attachment sites ($attB_C$, $attB_E$, $attB_B$ and $attB_F$ sites) were found to lie on a 200 kb segment of the chromosome,

near *terC*. The latter region is of ultimate importance for cellular processes that interact with the chromosome shape and its organization [42,43].

As also previously reported for other bacteria [41], prophage polarization was observed in *O. oeni*. The genes of oenophages show a preference for co-orientation with the bacterial replication fork. Accordingly, Int_D prophages, which are on the left replichore, changed their orientation.

Phages can alter host phenotype by disrupting host genes as a result of integration. Others are known to encode cis-acting element(s) near their *attP* site that, upon integration, affect the transcription of host neighboring genes [44,45]. We therefore scrutinized the annotated genomes for all nine integration loci in *O. oeni* and retrieved the genes surrounding the *attB* sites for functional classification (average distance was 4200 bp) (Figure 1). Integration occurred next to some genes which play important roles in bacterial cellular systems, and specify proteins involved in cell wall metabolism such as the WalK/WalR two-component system [46], response to osmotic stress (mechanosensitive channel of large conductance, MscL), a quality control system monitoring protein synthesis (SsrA) or carbohydrate metabolism (glucokinase, acetoin reductase, catabolite control A (CcpA)). CcpA is involved in the adaptation to changing carbon sources. More recently, its role as a pleiotropic regulator, controlling not only carbohydrate metabolism but also stress response, has been proposed in different LAB under aerobic conditions [47]. The hypothesis that some prophages/PRs may act as switches that regulate the expression of neighboring genes is attractive in *O. oeni* and deserves further work as this may have practical significance for the production of MLF starter cultures.

Of note, two sites ($attB_C$ and $attB_F$) could accommodate a PR, or a PR and a phage in a tandem association. We re-examined the sequences flanking these PRs and/or prophages amongst relevant strains and found that the two expected *attL* and *attR* junctions were conserved in tandem associations at the $attB_D$ site but not at the $attB_C$ locus. The latter consisted of a 78 bp repeated sequence, which may represent the ancestral $attB_C$ site. This size is larger than proposed earlier [15,17,38]. In all combinations (PR_C alone or tandem association with an Int_C prophage), we consistently identified a degenerated $attR_C$ sequence in the remnant, which may be indicative of its domestication. PR_C had maintained >99% nucleotide similarity across its entire genome, suggesting that it is under strong evolutionary pressure and likely provides an important biological function to the host. Its prevalence in the whole population (75.3%) was higher compared to Int_C phages (Table 1). It is likely that the acquisition of PR_C preceded that of Int_C phages during the evolution of *O. oeni*. The loss of the $attR_C$ site resulting from domestication did not hamper further integration of Int_C phages at the intact $attL_C$ site (Figure 2). In contrast, the ability of the intact Int_C prophages to excise is likely to be impaired as we did not observe any spontaneous or MC-induced excision of the phage in strains containing an Int_C phage/PR_C tandem association, such as S28.

3.4. Lysogeny Is Widespread among Wine Strains from Phylogroup A

We next explored whether quantitative and qualitative differences in patterns of prophages (as assessed from their integrase and attachment site used for SSR) differ across lineages and niches occupied by strains of *O. oeni* (Table 1). Most of the currently sequenced *O. oeni* strains in our set have, to date, been collected from wines and belong to phylogroup A (75%; n = 174). This enables a robust assessment of lysogeny in this particular phylogroup/niche.

Phylogroup A was particularly open to the uptake of foreign DNA of phage origin since 63% of the strains were lysogens (Table 1). Yet, prophages were not distributed uniformly across the different sub-groups described in the lineage. Interestingly, prevalence differed between the AR and AW sub-lineages, corresponding to strains adapted to red and white wines, respectively [10]. Hence, 100% of strains (n = 10) in sub-group AW associated with white wines produced in Burgundy and Champagne were lysogens [9]. In contrast, the proportion was only 40% (n = 10) in the Burgundy red wines subgroup (AR). Lysogeny

was less frequent in other sub-groups of strains in phylogroup A. As an example, only one of the seven strains constituting the PSU-1 subgroup was observed to carry a prophage [9], which is lower than the mean value observed in the whole *O. oeni* species.

Phylogroup A strains appear enriched in some prophages, corresponding to members of the Int_A, Int_B and Int_C groups. Conversely, and with few exceptions, these prophages were less abundant or absent in other lineages (phylogroups B, C and D) (Table 1). Of note, despite different origins of wine and time of collection, all 10 strains from sub-group AW (white wines) exclusively harbored an Int_A phage. Eight of ten had the same organization and sequence. This may reflect selection for the acquisition of locally adaptive functions that are transferred by the Int_A phage genomes. Alternatively, the Int_A phages may have piggybacked on hosts that outcompeted other variants in the process of natural selection in the specific white wines considered [48].

The second pattern seen was the frequent carriage of two distinct prophages in phylogroup A strains, as 45 out of the 47 poly-lysogens detected in the whole *O. oeni* species belonged to this particular phylogroup. This may suggest that interactions between prophages may be beneficial for the host, by reducing the rate of spontaneous lysis or regulating gene expression under specific wine conditions [27]. This hypothesis is attractive and needs more experimental support. Phylogroup A strains contained the six specific combinations of prophage carriage observed in the whole species: Int_A + Int_C (n = 17), Int_B + Int_C (n = 14), Int_A + Int_D (n = 8), Int_A + Int_B (n = 3), Int_A + Int_C + Int_D (n = 3) and Int_A + Int_B + Int_C (n = 1) (Table 1). Of note, prophages from the Int_D groups were found in 14 lysogens from phylogroup A, of which 11 corresponded to poly-lysogens. Since Int_D phages are the only prophages integrated on the left replichore, their presence and possibly low induction rate could balance the prophage integrations on the other replichore and stabilize the genome architecture. Although mono-lysogens for Int_D and Int_B prophages are found in wine strains, no Int_D + Int_B double lysogens were detected in phylogroup A, nor in the whole population. Each of these prophages may encode a resistance mechanism to phage superinfection. Alternately, the prophages may be incompatible with each other, or their presence may decrease cell fitness and lead to the extinction of the poly-lysogens in wine.

Further insight into the peculiarity of phylogroup A strains was also provided when we re-examined the distribution of PR elements (Table 1). We found that PR_C was present in all strains from phylogroup A, suggesting the domestication of the corresponding ancestral phage in wine strains, possibly leading to niche-specific fitness effects. These genetic signatures might also have applications as they could be used for typing purposes.

Phylogroups B, C and D are less numerously represented in our set of 231 strains (Table 1). Yet, with this caveat in mind, differences in prophage content, with respect to phylogroup A, were found. We detected prophages with lower frequency in phylogroup C (38%) while the value in phylogroup B (55%) was closer to that reported in phylogroup A (60%). Both phylogroups are known to contain strains originating from cider or wine, and the latter have been consistently isolated following completion of AF [9]. Patterns of prophages were also different in phylogroups B and C and Int_D phages were the most represented prophages. In particular, they corresponded to 12 out of the 15 prophages identified in mono-lysogens in phylogroup B. Only two poly-lysogens were found. Interestingly, phylogroup C was characterized by the presence of the specific Int_E prophages, observed in two cider strains and in strain AWRIB663 from wine. The same characteristic was observed for Int_F prophages, as well as PR_G, which have, to date, been exclusively found in kombucha strains. Due to the limited number of strains in the B, C and D phylogroups, further investigation and isolation campaigns on ciders, kombucha and possibly other fruits and kefirs are now needed.

3.5. Integrase Phylogeny

We performed a multiple alignment of phage-related integrases from *O. oeni*. We first included a representative sequence from each of the $Int_{A\text{-}F}$ phage groups and all four

PR-encoded integrases. The size of the 10 integrases ranged from 340 to 389 aa and we found 25 invariant amino acid residues, of which 10 are located in the C-terminal catalytic domain of the phage recombinases. Amongst them, the catalytic residue tetrad Arg–His–Arg–Tyr (R–H–R–Y), which is needed for DNA cleavage and joining in the integrase family of tyrosine recombinases, was identified (Figure S2).

In order to construct a maximum likelihood tree, we next included integrases from phages infecting the related species *O. sicerae* from ciders and kefir, *O. kitaharae* from sake [29–31] and other LAB belonging to *Pediococcus*, *Lactiplantibacillus*, *Fructilactobacillus* and *Streptococcus* genera [49] (Figure 3). The putative *O. oeni* XerC/XerD recombinases served as an outgroup. Of note, both XerC and XerD from *O. oeni* resemble the XerS protein which is involved in the cell recombination machinery of *Lc. lactis* [50].

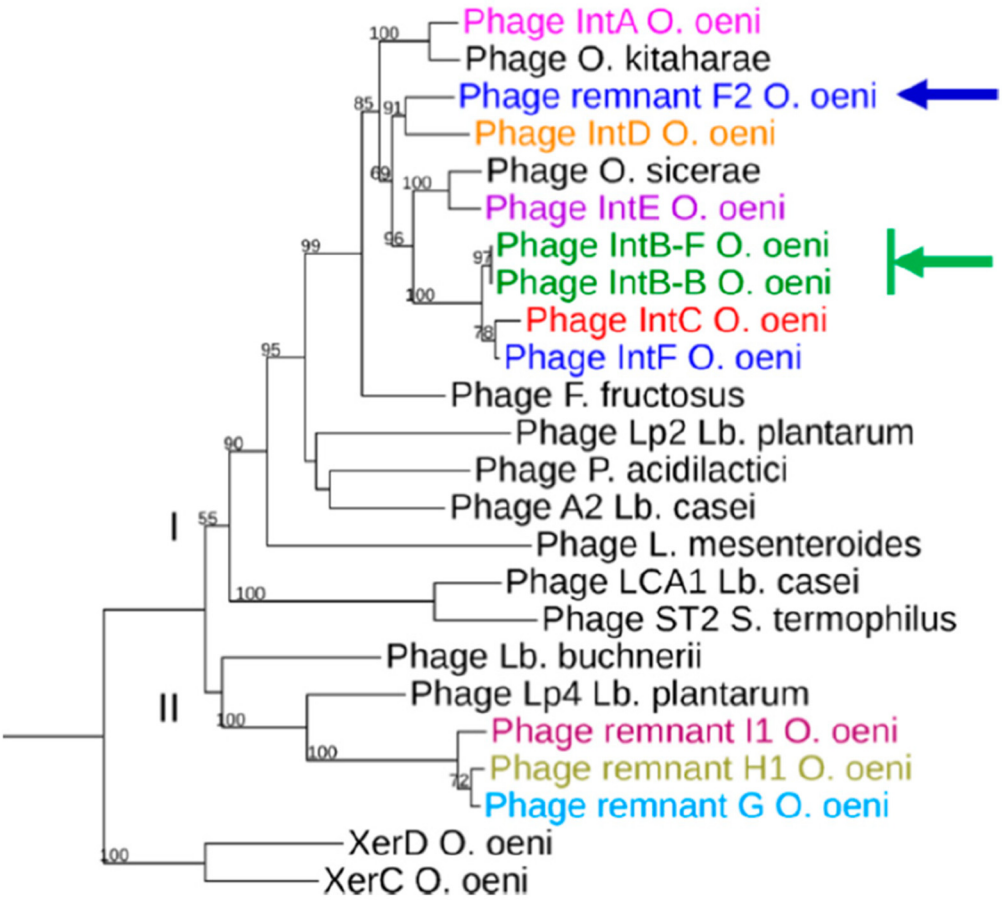

Figure 3. Phylogenetic tree of integrases found in *O. oeni* and other LAB. Sequences are distributed in clusters I and II. Their origin is as follows: Int_A (IOEB0608); Int_B Φ10MC (AAD00268); Int_C (IOEBS28); Int_D (IOEB9805); Int_E CRBO1384; Int_F (UBOCC315001); Int PRI_1 (IOEBS277); Int PRH_1 (AWRIB422); Int PR_G (UBOCC315001); Int PRF_1 (DIV5-23); Int from phages Lca1 (ABD83428) and A2 (NP_680502); unnamed phage infecting *P. acidilactici* 7-4 (ZP_06197614); Int from prophages found in *Lb. plantarum* WCFS1 (NP_785910), STIII (ADN97941) and Lp2 prophage (WP_011101785.1); *O. sicerae* (WP_128686695.1); *O. kitaharae* (EHN59022); *L. mesenteroides* ATCC19254 (ZP_03913849) and *Lb. buchneri* ATCC 11577 (ZP_03943244). The XerC and XerD proteins were obtained from *O. oeni* PSU-1 (WP_002817147 and WP_002820463.1) and served as an outgroup. Arrows represent incongruences between integrase phylogeny and *attB* location of the corresponding prophages in the bacterial genomes. The PRF_2-associated integrase-coding gene was split into two ORFs and carried a frameshift mutation due to a missing nucleotide. A single nucleotide was added in the premature stop codon to suppress the nonsense mutation.

As seen in Figure 3, most integrases identified in *O. oeni* belonged to two major clusters, suggesting distinct evolutionary trajectories for each group. The majority of phage integrases formed a cluster (cluster I) with sequences from phages infecting other LAB (*Lb. plantarum*, *P. acidilactici*, *Lb. casei* and *L. mesenteroides)*. This is consistent with the current literature showing the existence of cross-transmission networks between *O. oeni* and other LAB, as a variety of species coexist on grapes and must, including, notably, *Lactiplantibacillus plantarum* [13,49,51]. In addition, recent phylogenomics also resulted in the assignment of *Leuconostocaceae* and *Lactobacillaceae* into a single family [49]. A sub-group in cluster I was specific to prophages infecting species of the *Oenococcus* and *Fructilactobacillus* genera. In the latter genus, *F. fructosus* is the only species found in wine [52]. Interestingly, the Int_A type integrase was separate, and grouped evolutionarily with a sequence from *O. kitaharae*. Remarkably, the Int_E sequence grouped with the integrase found in a prophage from *O. sicereae.* The corresponding lysogens (affiliated to the *O. oeni* and *O. sicerae* species) were both isolated from cider. It would therefore be interesting to assess whether the oenophages can infect individual hosts from different species in cider and other fermented beverages. We also observed that the three integrases Int_B, Int_C and Int_F, which drive the integration of their cognate phages in the same region in the host chromosome, were more closely related (Figure 1). Lastly, the only integrase from a remnant in cluster I was that from PR_F. In contrast, the integrases harbored by $PR_{H1\text{-}H2}$, $PR_{I1\text{-}I2}$ and PR_G formed a distinct lineage, and clustered with two integrases associated with other LAB species. The inspection of the alignment showed that these five PR-associated integrases had, notably, an additional domain of 20 amino acids upstream of the tyrosine catalytic residue compared to other oenophage integrases. They may not represent remnants of previous lysogenization by full-length prophages, but rather belong to a unique family of mobile genetic elements.

3.6. Correlation of Phage Phylogeny with attB Location

The correlation of phage integrase phylogeny with *attB* location on the bacterial chromosome was confirmed with two notable exceptions, which were both related to the occupancy of the $attB_F$ site (Figures 1 and 3). This particular locus, close to the replication termini, could indeed accommodate an Int_B prophage, or an Int_F prophage and/or a PR_F remnant in the different strains of our panel. This raised two interesting questions. How did the Int_B recombinase display a relaxed specificity, driving the integration of the phage genome (hereby an Int_B group member) at distinct locations ($attB_B$ or $attB_F$) in the chromosome of distinct strains of *O. oeni*? Next, we also questioned the reason why PR_F integrates at the $attB_F$ site, although its cognate integrase was more closely related to members of the Int_D than to the Int_F group. The rationale behind these observations and questions was that a more detailed analysis of the SSR units of these oenophages and PRs was needed.

3.6.1. SSR of Int_F and Int_B Prophages

During SSR, cross-activity of a recombinase of one phage with the attachment site of another is proportional to the degree of homology between their integrases and similarity between the core- and arm-type sequences in attachment sites of the respective phages [53]. We therefore clarified the site preference pattern and sequence requirements for the SSR of Int_F and Int_B phages by analyzing their *attP* sites. To carry out the analyses, phage attachment core sequences were deduced from the *attL* and *attR* junctions within the host chromosome identified in lysogens (accessed from GenBank) and from the *attP* sequences from the genome of free Int_B phages isolated from the enological environment [18,54]. For clarity, the phages belonging to the Int_B group were subdivided into the $Int_{B\text{-}F}$ and $Int_{B\text{-}B}$ sub-groups, depending on the attachment site targeted in the bacterial chromosome ($attB_B$ or $attB_F$). Prophages harbored by strain LAB2013 and IOEBB10 were their representatives, respectively. Int_F phages were represented by the prophage of strain UBOCC315001.

SSR of Int_F phages was found to complement the 13 bp 3′ end of the bacterial $tRNA^{Leu\ 923}$. This sequence does not encompass the tRNA anticodon loop. The iden-

tity block common to the phage and bacterial sequences extended well beyond, and a 63 bp homologous pair was identified (Figure 4A). It is well established that SSR requires a longer sequence at *attP* than at *attB*. Tyrosine recombinases usually possess a low binding affinity to the core site of *attP* and a high binding affinity for the flanking arm regions [53]. We next questioned what sequences within $attP_F$ actually support integrase binding in *O. oeni*. We addressed this question by screening the arm regions for direct repeats (DRs), which could serve as putative arm-binding sites recognized by the N-terminus of the integrase (Figure 4A). A short 10 bp DR was found (5′ATTTGCACAA3′, Figure 4A). It was not evenly distributed on both arms around the core, with two and four repeats on the left and right arms (P1 to P6), respectively (Figure 4A and Table 3). Like other characterized systems amongst LAB phages, the proposed core site exhibited essentially no symmetry and no putative inverted core-binding sites were identified [55]. With the caveat that additional putative binding sites for accessory proteins (excisionase, integration host factor) may be present, we propose that the size of the *attP* sequence of Int_F phages is at least 240 bp. This size is consistent with the well-documented model proposed by Campbell for λ integration in *Escherichia coli* [56,57].

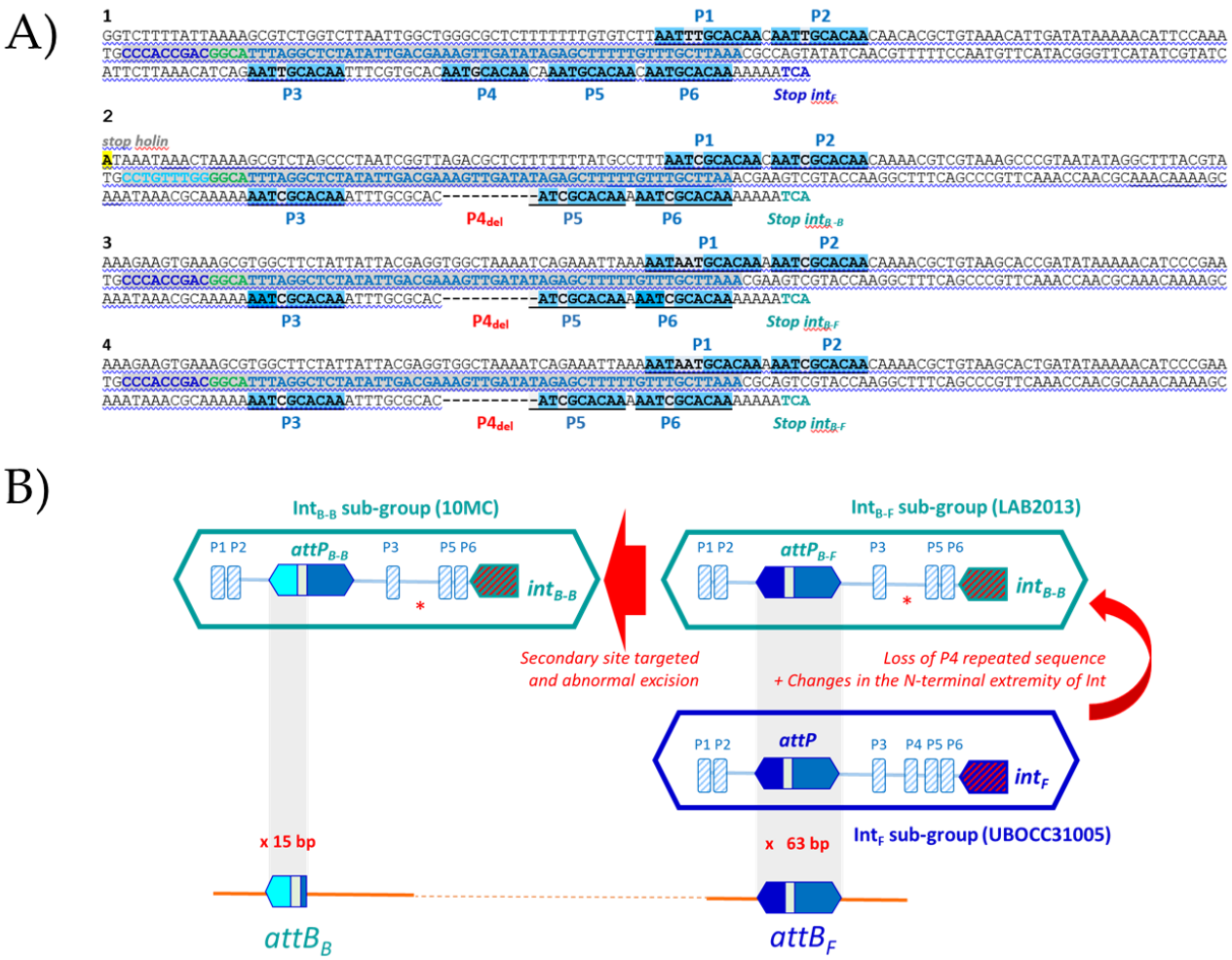

Figure 4. Interrogation of oenococcal genomes for the SSR units of Int_{B-B}, Int_{B-F} and Int_F prophages. (**A**) Alignments of the *attP* regions involved in SSR for Int_F phages (1 in blue), the proposed sub-types Int_{B-B} and Int_{B-F} phages (3 and 4 in green). 1, Int_F prophage from strain UBOCC315005; 2, Int_{B-B} prophage from strain IOEBB10; 3, Int_{B-F} prophage from strain LAB2013; 4, Int_{B-F} phage OE33PA. P1 to P6, Direct repeats. The 63 bp block identity is in gray. (**B**) Boxed sequences represent the *attP* sequences of Int_B and Int_F phages. Arrows indicate the possible complex mechanisms of evolution of SSR units during niche adaptation.

Recombination of $Int_{B\text{-}F}$ prophages also required the same 63 bp core sequence, common to *attP* and $attB_F$. However, the detailed analysis of the whole $attP_{B\text{-}F}$ region showed some deletions, resulting in the loss of the P4 repeat on the right arm (Figure 4A, Table 3). As a consequence, there was a reduction in the number of DRs, as well as a modified spacing between the P3 and P5 repeats, which may impact protein–protein and/or protein–DNA interactions during intasome formation. Another striking difference was found when the $Int_{B\text{-}F}$ and Int_F protein sequences were compared (Figure S2). Despite their high identity (92%), they differed by 46 amino acid (aa) changes, of which 21 (45%) mapped to the 66 aa amino terminus, representing 18% of the integrase. Similar results were observed with the Int_B integrases obtained from the free replicating phages OE33PA and B148 previously isolated from wine [17,18]. The size of the Int_F N-terminus domain which concentrates the mutations (66 aa) is in agreement with the assignment of the minimal Int fragment of 64 aa that binds to arm-type sites in the lambda model [58].

Table 3. Characteristics of the *attB* and *attP* regions from oenophages and phage remnants.

Phage-Related Elements			Host Strain and *attB* Site		Putative Integrase Binding Sites (DR *) on *attP*		
P/PR	**Strain Harboring the Prophage or PR Name**	**Int**	**Name, Phylogroup and Niche**	***attB* Site (bp)**	**nb**	**Right/Left Arms**	**DR Sequences**
P	IOEB0608	Int_A	IOEB0608 (A/wine)	$attB_A$ ($tRNA^{Glu}$, 16 bp)	5	2/3	-TTT-GCCACA (x2) -TTT-GCCACA (x2) -TTT-GCCACA (x1)
	10MC	$Int_{B\text{-}B}$	IOEB B10 (A/wine)	$attB_B$ ($tRNA^{Leu}$, 15 bp)	5	2/3	–ATCG-CACAA (x5) –ATCG-CACAA (x1)
	LAB2013	$Int_{B\text{-}F}$	LAB2013 (A/wine)	$attB_F$ ($tRNA^{Leu}$, 63 bp)	5	2/3	–AT-G-CACAA (x1) –ATCG-CACAA (x3) –ATCG-CACAA (x1)
	IOEBS28	Int_C	IOEB S28	$attB_C$ ($tRNA^{Lys}$, 78 bp)	4	2/2	–AT-G-CACAACAA (x3) –CT-G-CACAACAA (x1)
	IOEB9805	Int_D	IOEB9805 (B/wine)	$attB_D$ ($tRNA^{Leu}$, 128 bp)	5	2/3	-TTT-G-CACA (x5)
	CRBO1384	Int_E	CRBO1384	$attB_E$ ($tRNA^{Cyst}$, 11 bp)	4	2/2	T-GCCAC-CGTT (x4)
	UBOCC315005	Int_F	UBOCC315005 (D/kombucha)	$attB_F$ ($tRNA^{Leu}$, 63 bp)	6	2/4	ATTT-G-CACAA (x1) AT-T-G-CACAA (x2) AA-T-G-CACAA (x3)
PR	PR_C	Int_C	IOEB S28	$attB_C$ ($tRNA^{Lys}$, 78 bp)	4	2/2	–AT-G-CACAACAA(x3) –CT-G-CACAACAA (x1)
	PR_{F1}	Int_{F1}	DIV5-23 (B/cider)	$attB_F$ ($tRNA^{Leu}$, 63 bp)	5	3/2	—TGG-TACAA (x4) —TGGCTACAA (x1)
	PR_{F2}	Int_{F2}	CRB01381 (C/cider)	$attB_F$ ($tRNA^{Leu}$, 63 bp)	6	3/3	—TGG-TACAA (x4) —TAGCTACAA (x1) —TGG-TCAAA (x1)
	PR_G	Int_G	UBOCC315005	$attB_G$ ($tRNA^{Cyst}$, 61 bp)	nf		
	PR_{H1}	Int_{H1}	AWRIB422	$attB_H$ ($tRNA^{Ser}$, 22 bp)	nf		
	PR_{H2}	Int_{H2}	IOEBC52	$attB_H$ ($tRNA^{Ser}$, 22 bp)	nf		
	PR_{I1}	Int_{I1}	IOEB277	$attB_I$ ($tRNA^{Ser}$, 19 bp)	nf		
	PR_{I2}	Int_{I2}	AWRIB576	$attB_I$ ($tRNA^{Ser}$, 19 bp)	nf		

* DR, Direct repeat; nf, not found.

Altogether, our data are consistent with the existence of different N-termini in Int_F and $Int_{B\text{-}F}$ integrases, allowing each recombinase to bind the specific sets of DRs on the arm regions in each *attP* site. The observed mutations may have a broader impact on SSR since the N-terminal domain of Int also plays a role in modulating the activity of core-binding and catalytic domains [59]. In addition, the impact of the mutations associated with the other domains of the integrase protein sequences also needs further work.

Noteworthily, Int_F and $Int_{B\text{-}F}$ phages were associated with strains from distinct niches, corresponding to kombucha and wine, respectively. It is likely that the presence of slight modifications in their SSR units needs to be examined in light of the distinct niches where

their bacterial hosts are evolving. Our observations may reflect the coevolution of the $attP_F$ and integrase sequences in response to changing conditions. They may be a signature of distinct evolutionary trajectories across the different niches colonized by *O. oeni*, and represent signatures of adaptation to wine. We posit that modifications in the $attP_F$ site may have progressively arisen during the adaptation of *O. oeni* to more drastic variations in habitat conditions (wine) and have been coupled to mutations in the Int_F sequence. The latter adapted the novel integrase (Int_B) to this new and altered site, resulting in the variations observed within the SSR units of Int_F and $Int_{B\text{-}F}$ phages [60]. Altogether, our observations may reflect the evolutionary processes resulting in niche divergence. If this hypothesis is correct, it questions the existence of selection for the preservation of $attB_F$ as an integration site for phages in two distinct niches (kombucha and wine).

Last, it is likely that $Int_{B\text{-}F}$ phages went through additional and distinct evolutionary processes in wine, yielding $Int_{B\text{-}B}$ phages. Hence, the latter prophages were shown to integrate into a distinct gene, 66 kb away from the $attB_F$ site used by $Int_{B\text{-}F}$. Both sites were part of $tRNA^{Leu}$ genes. The SSR of $Int_{B\text{-}F}$ and $Int_{B\text{-}B}$ phages required the same overall region in the phage genomes, but the nature of the sequences involved were different. Striking differences found in $Int_{B\text{-}B}$ phages included a reduced number of DRs on arms (5) and the substitution of the 5′ extremity of the 63 bp sequence by a new 9 bp sequence. In addition, the identity block between host and $Int_{B\text{-}B}$ phage sequences was reduced to the first 15 bp of the previously proposed 63 bp core (Figure 4B). Since the Int_B integrase recombines regardless of the lateral 9 bp sequence in the *attP* site present in $Int_{B\text{-}B}$ and $Int_{B\text{-}F}$ phages, the latter nucleotides probably do not represent core-binding sequences for the integrase. The role is probably devoted to the downstream 6 bp sequence that is common to $attP_{B\text{-}B}$ and $attP_{B\text{-}F}$ sites and where cleavage is likely to occur. The emergence of $Int_{B\text{-}B}$ phages is likely to result from mutations leading to a much more substantial increase in the recognition of a secondary site by $Int_{B\text{-}F}$ phages followed by an abnormal excision. This hypothesis is consistent with the chromosome jumping model described in lambda [61].

3.6.2. SRR of PR_F Elements

The PR_{F1} and PR_{F2} remnants are integrated in the $attB_F$ site in strains collected from cider and wine and from kombucha (phylogroups B and D, respectively). Yet, their integrase was more related to the Int_D than to the Int_F type (Figure 3). Both integrase-coding gene were split into two ORFs in PR_{F1} and PR_{F2}, possibly due to a frameshift mutation (Figure 2). The sequence was manually inspected and a single nucleotide was added in the premature stop codon to suppress the nonsense mutation and the corrected deduced protein sequence was used to build the phylogeny of integrase proteins (Figure 3). We assessed whether an error at this step may have caused an alignment bias, explaining the incongruence of the PR_F recombinase in the tree. To verify this, we compared the three *int* nucleotide sequences found in PR_F, Int_F and Int_D prophages (Figure S3). This confirmed the closer relation of the integrase nucleotide sequences from PR_F and Int_D prophages. The reason for the discrepancy between the integrase type and integration site of PR_F is not known. However, it can be suggested that PR_F elements have resulted from modular exchanges in the lysogeny modules between two integrated Int_D and Int_F oenophages, due to intra-chromosomal homologous recombination. Homologous recombination between two prophages integrated equidistant from the *ter* region has been reported in *S. pyogenes*. Such phage-related rearrangements resulted in a large chromosomal inversion of the region between the attachment sites, and the emergence of two novel hybrid prophages with exchanged genes [62]. In *O. oeni*, int_D and int_F sequences contain homologous regions. In addition, such homologous recombination events may also be mediated by the bacterial sequences flanking each prophage which both correspond to a $tRNA^{Leu}$ gene.

3.7. Most attP Regions in Oenophages Derive from Two Distinct Sequences

Surprisingly, all *attP* sequences upstream of the integrase gene in Int_A, Int_B, Int_C, Int_D and Int_F prophages had similar sequences, suggesting a common origin (Figure 5).

Despite the presence of several indel events, all core sequences involved in RSS (underlined in Figure 5) had a common feature and were flanked by direct repeat sequences, which are proposed to represent the binding sites for the different recombinases (Table 3). Coevolution of the phage components of SSR units have progressively occurred and involved indels in the core, the slippering of DR sequences and modifications in the N-terminus of the integrase (Figure 5). Through such events, recombinase activity is retained while the core *attP* sites are progressively adjusted to novel loci in the chromosome of the host. This may indicate long-term coexistence between these phages and the host in fermented beverages.

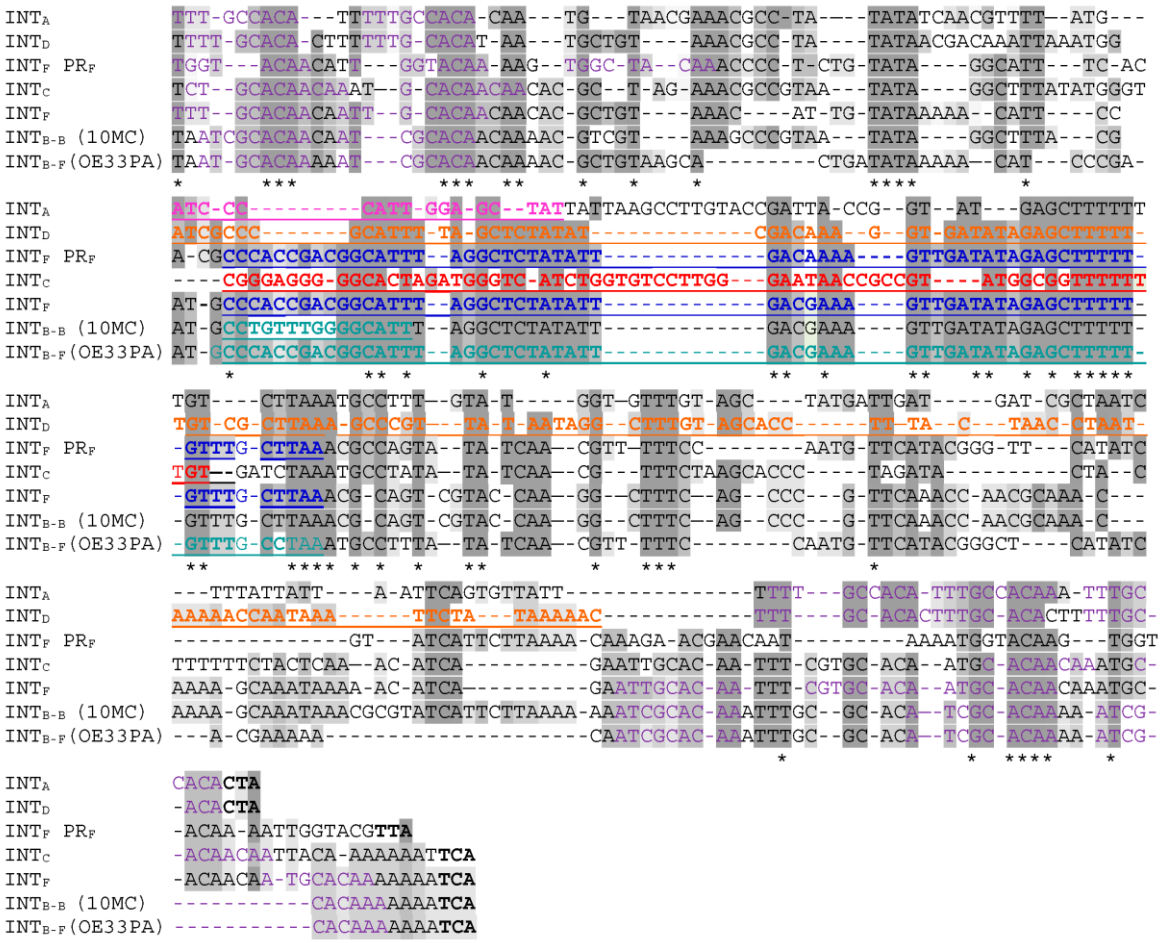

Figure 5. Alignments of the *attP* sequences found in the different oenophages. The homologous core sequences involved in recombination with the bacterial chromosome are underlined and in color (the color code is the same as shown in Figure 1). All direct repeats serving as putative binding sites for integrases are in purple. The three last nucleotides correspond to the stop codon of the integrase genes. A clustal alignment was first constructed and gaps were introduced manually.

Of note, the *attP*s from Int_E phages were distinct in their sequence and also lacked DR repeats compared to other oenophages. Future studies should experimentally confirm whether the Int_E recombinase can utilize the candidate core sequences for recombination reactions without flanking inverted repeats. The comparative analysis of a set of completely sequenced oenophage genomes has recently demonstrated that prophages are distributed into two clusters of *cos* (members of Int_A and Int_B groups) and *pac* (Int_D) phages [63]. Even though the SSR unit of Int_E phages has some peculiarities, the latter have not evolved completely independently. As seen in Figure 6, Int_E phages are closely related to the *cos* phages, together with Int_C and Int_F prophages. Corresponding genomes in this cluster are mosaics, whereby individual phages are constructed as assemblages of modules, many of which are single genes (result not shown). Of note, this cluster harbors prophages which

are integrated on the same replichore in the chromosome of *O. oeni*. In contrast, the cluster corresponding to Int_D prophages is less diverse and members share relatively few genes with the other cluster.

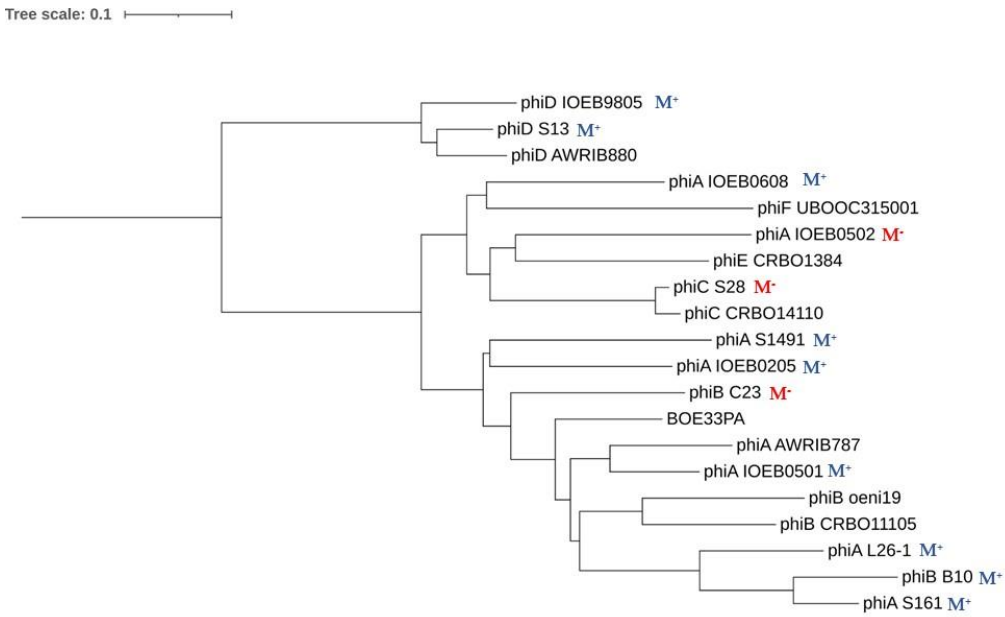

Figure 6. Phylogenomic Genome BLAST Distance Phylogeny (GBDP) tree of representative oenophages from the CRB Oeno Collection. The tree was generated by VICTOR and visualized with Fig Tree [36]. Int groups are represented by letters (A to F). B10 is the prophage found in strain IOEB B10. M^+ and M^- indicate the presence or absence of (i) a lysis when the lysogen was grown in the presence of mitomycin C and (ii) plaques when the MC lysate was tested on the sensitive strain IOEB S77.

4. Concluding Remarks

A short minimal doubling time under optimal growth conditions was shown to represent the trait most correlated with lysogeny among bacteria, and the frequency of lysogeny was also increased with bacterial genome size [41]. Clearly, *O. oeni* stands out from the well-studied model systems used in this study, and probably faces specific ecological conditions which constrain the lytic–lysogeny decision and favor lysogeny and poly-lysogeny. This may generate more benefits than costs for both partners.

Our results also confirmed that tRNA genes are the preferred chromosomal integration sites in *O. oeni* and that integration tropisms are associated with the phylogeny of the phage integrases. Interestingly, all bacterial attachment sites, except $attB_D$, are located in one replichore (one half of the chromosome) and four lie within a 300 kb region, which is therefore a region of high plasticity in the species. Such lopsided phage integrations into chromosomal DNA may result in an unsymmetrical genome architecture across the replication axis, and induce chromosomal rearrangement for stabilizing the genome architecture. Yet, such rearrangements are not observed in *O. oeni* [9].

The different niches where the *Oenococcus* species is found have unique physical, chemical and biological profiles that likely promote speciation of both phages and their bacterial hosts. Accordingly, phages infecting strains associated with kombucha, cider and wines were observed to exhibit differences in their SSR units and more work is now needed to explore their gene repertoire as well, and assess whether they provide the bacterial hosts with additional genes and competitive advantages.

Considering whether prophages (or combinations thereof) have a positive role in the adaptation of *O. oeni* to wine, it could be expected that strains would domesticate such beneficial prophages, as seen with PR regions, in order to prevent their excision. This is probably the case for Int_C prophages, as we could not yet detect excision of phage DNA/particles in past studies. In contrast, many pieces of experimental evidence show

that members of the frequently encountered Int_A and Int_B prophages in wine are still active and can excise. Therefore, it is unclear why Int_A and Int_B phages are less prone to grounding, compared to members of the Int_C group. Reversible lysogeny could be of importance for genome architecture, as the grounding of the various prophages located on the right replichore could be detrimental. In addition, the ability of such prophages to excise could represent a competitive mechanism to eliminate sensitive non lysogenic strains, or to lysogenize them, increasing the tolerance of the population to stressful conditions. Conversely, site-specific integration of Int_A and Int_B phages may be beneficial under certain circumstances and modulate the expression of essential genes in stressful conditions. These characteristics might reflect the different evolutionary strategies and opposite selection pressures as a consequence of adaptation to diverse niches in which the different phylogroups have evolved.

Supplementary Materials: The following materials are available online at https://www.mdpi.com/article/10.3390/microorganisms9040856/s1. Table S1: Genomes of lysogenic and non-lysogenic strains of *O. oeni* used in this study; Figure S1: Kinetics of bacteriophage induction from the lysogen *O. oeni* IOEB0608 with (■) and without (□) mitomycin C treatment (1 µg/mL). Each culture was grown in MRS broth to an OD_{600} of 0.2–0.3 and then separated into two aliquots, one of which was induced with mitomycin C; Figure S2: Comparisons of the Int_F and Int_B protein sequences. Mutations are in gray. The four conserved amino acids in the C-terminal region (catalytic domain) of all integrases found in *O. oeni* are in red; Figure S3: Alignments of the nucleotide sequences of the integrase genes from PR_F and Int_F and Int_D prophages. The yellow color represents deletions in $intPR_F$ and int_D sequences. The blue color represents deletions only observed in the int_F sequence.

Author Contributions: O.C., C.L.M., P.M.L. designed experiments, assembled and analyzed the genomic data and prepared the manuscript. A.C., F.J., C.P., Y.B. performed experimental work. C.L.M. supervised the work progress and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: Support for this project was provided by the Conseil Regional Nouvelle Aquitaine and French ANR (grant ANR-2020-CE20-0008-01).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Grandvalet, C. *Oenococcus oeni*: Queen of the cellar, nightmare of geneticists. *Microbiology* **2017**, *163*, 297–299. [CrossRef] [PubMed]
2. Lorentzen, M.P.G.; Lucas, P.M. Distribution of *Oenococcus oeni* populations in natural habitats. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 2937–2945. [CrossRef] [PubMed]
3. Mills, D.A.; Rawsthorne, H.; Parker, C.; Tamir, D.; Makarova, K. Genomic analysis of *Oenococcus oeni* PSU-1 and its relevance to winemaking. *FEMS Microbiol. Rev.* **2005**, *29*, 465–475. [CrossRef]
4. Borneman, A.R.; McCarthy, J.M.; Chambers, P.J.; Bartowsky, E.J. Comparative analysis of the *Oenococcus oeni* pan genome reveals genetic diversity in industrially-relevant pathways. *BMC Genom.* **2012**, *13*, 373. [CrossRef]
5. Dimopoulou, M.; Vuillemin, M.; Campbell-Sills, H.; Lucas, P.M.; Ballestra, P.; Miot-Sertier, C.; Favier, M.; Coulon, J.; Moine, V.; Doco, T.; et al. Exopolysaccharide (EPS) synthesis by *Oenococcus oeni*: From genes to phenotypes. *PLoS ONE* **2014**, *9*, e98898. [CrossRef]
6. Sternes, P.R.; Borneman, A.R. Consensus Pan-Genome Assembly of the Specialised Wine Bacterium *Oenococcus oeni*. *BMC Genom.* **2016**, *17*, 813.
7. Campbell-Sills, H.; Khoury, M.E.; Favier, M.; Romano, A.; Biasioli, F.; Spano, G.; Sherman, D.J.; Bouchez, O.; Coton, E.; Coton, M.; et al. Phylogenomic analysis of *Oenococcus oeni* reveals specific domestication of strains to cider and wines. *Genome Biol. Evol.* **2015**, *7*, 1506–1518. [CrossRef] [PubMed]
8. El Khoury, M.; Campbell-Sills, H.; Salin, F.; Guichoux, E.; Claisse, O.; Lucas, P.M. Biogeography of *Oenococcus oeni* Reveals Distinctive but Nonspecific Populations in Wine-Producing Regions. *Appl. Environ. Microbiol.* **2017**, *83*, e02322-16. [CrossRef]
9. Lorentzen, M.P.; Campbell-Sills, H.; Jorgensen, T.S.; Nielsen, T.K.; Coton, M.; Coton, E.; Hansen, L.; Lucas, P.M. Expanding the biodiversity of *Oenococcus oeni* through comparative genomics of apple cider and kombucha strains. *BMC Genom.* **2019**, *20*, 330. [CrossRef] [PubMed]

10. Campbell-Sills, H.; El Khoury, M.; Gammacurta, M.; Miot-Sertier, C.; Dutilh, L.; Vestner, J.; Capozzi, V.; Sherman, D.; Hubert, C.; Claisse, O.; et al. Two different *Oenococcus oeni* lineages are associated to either red or white wines in Burgundy: Genomics and metabolomics insights. *OENO One* **2017**, *51*, 309. [CrossRef]
11. Marcobal, A.M.; Sela, D.A.; Wolf, Y.I.; Makarova, K.S.; Mills, D.A. Role of hypermutability in the evolution of the genus *Oenococcus*. *J. Bacteriol.* **2008**, *190*, 564–570. [CrossRef]
12. Meier, P.; Wackernagel, W. Impact of *mutS* inactivation on foreign DNA acquisition by natural transformation in *Pseudomonas stutzeri*. *J. Bacteriol.* **2005**, *187*, 143–154. [CrossRef]
13. El Gharniti, F.; Dols-Lafargue, M.; Bon, E.; Claisse, O.; Miot-Sertier, C.; Lonvaud, A.; Le Marrec, C. IS30 elements are mediators of genetic diversity in *Oenococcus oeni*. *Int. J. Food Microbiol.* **2012**, *158*, 14–22. [CrossRef] [PubMed]
14. Favier, M.; Bilhère, E.; Lonvaud-Funel, A.; Moine, V.; Lucas, P.M. Identification of pOENI-1 and related plasmids in *Oenococcus oeni* strains performing the malolactic fermentation in wine. *PLoS ONE* **2012**, *7*, e49082. [CrossRef] [PubMed]
15. Bon, E.; Delaherche, A.; Bilhère, E.; De Daruvar, A.; Lonvaud-Funel, A.; Le Marrec, C. *Oenococcus oeni* genome plasticity is associated with fitness. *Appl. Environ. Microbiol.* **2009**, *75*, 2079–2090. [CrossRef]
16. Poblet-Icart, M.; Bordons, A.; Lonvaud-Funel, A. Lysogeny of *Oenococcus oeni* (syn. *Leuconostoc oenos*) and study of their induced bacteriophages. *Curr. Microbiol.* **1998**, *36*, 365–369. [CrossRef]
17. Jaomanjaka, F.; Ballestra, P.; Dols-Lafargue, M.; Le Marrec, C. Expanding the diversity of oenococcal bacteriophages: Insights into a novel group based on the integrase sequence. *Int. J. Food Microbiol.* **2013**, *166*, 331–340. [CrossRef]
18. Philippe, C.; Jaomanjaka, F.; Claisse, O.; Laforgue, R.; Maupeu, J.; Petrel, M.; Le Marrec, C. A survey of oenophages during wine making reveals a novel group with unusual genomic characteristics. *Int. J. Food Microbiol.* **2017**, *257*, 138–147. [CrossRef] [PubMed]
19. Howard-Varona, C.; Hargreaves, K.R.; Abedon, S.T.; Sullivan, M.B. Lysogeny in nature: Mechanisms, impact and ecology of temperate phages. *ISME J.* **2017**, *11*, 1511–1520. [CrossRef]
20. Wahl, A.; Battesti, A.; Ansaldi, M. Prophages in Salmonella enterica: A driving force in reshaping the genome and physiology of their bacterial host? *Mol. Microbiol.* **2019**, *111*, 303–316. [CrossRef]
21. Ruiz-Cruz, S.; Parlindungan, E.; Erazo Garzon, A.; Alqarni, M.; Lugli, G.A.; Ventura, M.; van Sinderen, D.; Mahony, J. Lysogenization of a Lactococcal Host with Three Distinct Temperate Phages Provides Homologous and Heterologous Phage Resistance. *Microorganisms* **2020**, *8*, 1685. [CrossRef]
22. Matos, R.C.; Lapaque, N.; Rigottier-Gois, L.; Debarbieux, L.; Meylheuc, T.; Gonzalez-Zorn, B.; Repoila, F.; Lopes, M.d.F.; Serror, P. *Enterococcus faecalis* prophage dynamics and contributions to pathogenic traits. *PLoS Genet.* **2013**, *9*, e1003539. [CrossRef]
23. Aucouturier, A.; Chain, F.; Langella, P.; Bidnenko, E. Characterization of a Prophage-Free Derivative Strain of *Lactococcus lactis* ssp. lactis IL1403 Reveals the Importance of Prophages for Phenotypic Plasticity of the Host. *Front. Microbiol.* **2018**, *9*, 2032. [CrossRef]
24. Lynn-Bell, N.L.; Strand, M.R.; Oliver, K.M. Bacteriophage acquisition restores protective mutualism. *Microbiology* **2019**, *165*, 985–989. [CrossRef] [PubMed]
25. De Sordi, L.; Lourenço, M.; Debarbieux, L. "I will survive": A tale of bacteriophage-bacteria coevolution in the gut. *Gut Microbes* **2018**, *10*, 1–8. [CrossRef]
26. Braga, L.P.P.; Soucy, S.M.; Amgarten, D.E.; da Silva, A.M.; Setubal, J.C. Bacterial Diversification in the light of the interactions with phages: The genetic symbionts and their role in ecological speciation. *Front. Ecol. Evol.* **2018**, *6*, 6. [CrossRef]
27. Szafrański, S.P.; Kilian, M.; Yang, I.; Der Wieden, G.B.; Winkel, A.; Hegermann, J.; Stiesch, M. Diversity patterns of bacteriophages infecting *Aggregatibacter* and *Haemophilus* species across clades and niches. *ISME J.* **2019**, *13*, 2500–2522. [CrossRef] [PubMed]
28. Badotti, F.; Moreira, A.P.B.; Tonon, L.A.C.; De Lucena, B.T.L.; Gomes, F.D.C.O.; Krüger, R.; Thompson, C.C.; De Morais, M.A.; Rosa, C.A.; Thompson, F.L. *Oenococcus alcoholitolerans* sp. nov., a lactic acid bacteria isolated from cachaça and ethanol fermentation processes. *Antonie Van Leeuwenhoek* **2004**, *106*, 1259–1267. [CrossRef]
29. Endo, A.; Okada, S. *Oenococcus kitaharae* sp. nov., a non-acidophilic and non-malolactic-fermenting *Oenococcus* isolated from a composting distilled shochu residue. *Int. J. Syst. Evol. Microbiol.* **2006**, *56*, 2345–2348. [CrossRef]
30. Cousin, F.J.; Le Guellec, R.; Chagnot, C.; Goux, D.; Dalmasso, M.; Laplace, J.M.; Cretenet, M. *Oenococcus sicerae* sp. nov., isolated from French cider. *Syst. Appl. Microbiol.* **2019**, *42*, 302–308. [CrossRef]
31. Verce, M.; De Vuyst, L.; Weckx, S. The metagenome-assembled genome of Candidatus *Oenococcus aquikefiri* from water kefir represents the species *Oenococcus sicerae*. *Food Microbiol.* **2020**, *88*, 103402. [CrossRef]
32. McNair, K.; Aziz, R.K.; Pusch, G.D.; Overbeek, R.; Dutilh, B.E.; Edwards, R. Phage Genome Annotation Using the RAST Pipeline. *Methods Mol. Biol.* **2018**, *1681*, 231–238.
33. Czajkowski, R. May the Phage be With You? Prophage-Like Elements in the Genomes of Soft Rot *Pectobacteriaceae*: *Pectobacterium spp.* and *Dickeya spp*. *Front. Microbiol.* **2019**, *10*, 138. [CrossRef]
34. Arndt, D.; Grant, J.; Marcu, A.; Sajed, T.; Pon, A.; Liang, Y.; Wishart, D.S. PHASTER: A better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* **2016**, *44*, W16–W21. [CrossRef] [PubMed]
35. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [CrossRef] [PubMed]
36. Meier-Kolthoff, J.P.; Göker, M. VICTOR: Genome-based Phylogeny and Classification of Prokaryotic Viruses. *Bioinformatics* **2017**, *33*, 3396–3404. [CrossRef]

37. Philippe, C.; Krupovic, M.; Jaomanjaka, F.; Claisse, O.; Petrel, M.; Le Marrec, C. Bacteriophage GC1, a novel Tectivirus Infecting *Gluconobacter cerinus*, an acetic acid bacterium associated with wine-making. *Viruses* **2018**, *10*, 39. [CrossRef] [PubMed]
38. São-José, C.; Santos, S.; Nascimento, J.; Brito-Madurro, A.G.; Parreira, R.; Santos, M.A. Diversity in the lysis-integration region of oenophage genomes and evidence for multiple tRNA loci, as targets for prophage integration in *Oenococcus oeni*. *Virology* **2004**, *325*, 82–95. [CrossRef]
39. Petersen, A.; Josephsen, J.; Johnsen, M.G. TPW22, a lactococcal temperate phage with a site-specific integrase closely related to *Streptococcus thermophilus* phage integrases. *J. Bacteriol.* **1999**, *181*, 7034–7042. [CrossRef]
40. Van der Ploeg, J.R. Characterization of *Streptococcus gordonii* Prophage PH15: Complete Genome Sequence and Functional Analysis of Phage-Encoded Integrase and Endolysin. *Microbiology* **2008**, *154*, 2970–2978. [CrossRef]
41. Bobay, L.M.; Rocha, P.C.; Touchon, M. The Adaptation of Temperate Bacteriophages to Their Host Genomes. *Mol. Biol. Evol.* **2013**, *30*, 737–751. [CrossRef] [PubMed]
42. Touchon, M.; Rocha, E.P. Coevolution of the Organization and Structure of Prokaryotic Genomes. *Cold Spring Harb. Perspect. Biol.* **2016**, *8*, a018168. [CrossRef]
43. Kopejtka, K.; Lin, Y.; Jakubovičová, M.; Koblížek, M.; Tomasch, J. Clustered Core- and Pan-Genome Content on *Rhodobacteraceae* Chromosomes. *Genome Biol. Evol.* **2019**, *11*, 2208–2217. [CrossRef] [PubMed]
44. Feiner, R.; Argov, T.; Rabinovich, L.; Sigal, N.; Borovok, I.; Herskovits, A.A. A new perspective on lysogeny: Prophages as active regulatory switches of bacteria. *Nat. Rev. Microbiol.* **2015**, *13*, 641–650. [CrossRef] [PubMed]
45. Carey, J.N.; Mettert, E.L.; Fishman-Engel, D.R.; Roggiani, M.; Kiley, P.J.; Goulian, M. Phage integration alters the respiratory strategy of its host. *eLife* **2019**, *8*, e49081. [CrossRef]
46. Takada, H.; Yoshikawa, H. Essentiality and function of WalK/WalR two-component system: The past, present, and future of research. *Biosci. Biotechnol. Biochem.* **2018**, *82*, 741–751. [CrossRef]
47. Li, C.; Sun, J.W.; Zhang, G.F.; Liu, L.B. Effect of the absence of the CcpA gene on growth, metabolic production, and stress tolerance in *Lactobacillus delbrueckii ssp. bulgaricus*. *J. Dairy Sci.* **2016**, *99*, 104–111. [CrossRef]
48. Knowles, B.; Silveira, C.B.; Bailey, B.A.; Barott, K.; Cantu, V.A.; Cobián-Güemes, A.G.; Coutinho, F.H.; Dinsdale, E.A.; Felts, B.; Furby, K.A.; et al. Lytic to temperate switching of viral communities. *Nature* **2016**, *531*, 466–470. [CrossRef] [PubMed]
49. Zheng, J.; Wittouck, S.; Salvetti, E.; Franz, C.M.A.P.; Harris, H.M.B.; Mattarelli, P.; O'Toole, P.W.; Pot, B.; Vandamme, P.; Walter, J.; et al. A taxonomic note on the genus *Lactobacillus:* Description of 23 novel genera, emended description of the genus *Lactobacillus* Beijerinck 1901, and union of *Lactobacillaceae* and *Leuconostocaceae*. *Int. J. Syst. Evol. Microbiol.* **2020**, *70*, 2782–2858. [CrossRef]
50. Le Bourgeois, P.; Bugarel, M.; Campo, N.; Daveran-Mingot, M.L.; Labonté, J.; Lanfranchi, D.; Lautier, T.; Pagès, C.; Ritzenthaler, P. The unconventional Xer recombination machinery of Streptococci/Lactococci. *PLoS Genet.* **2007**, *3*, e117. [CrossRef]
51. Zúñiga, M.; Pardo, I.; Ferrer, S. Transposons Tn916 and Tn925 can transfer from *Enterococcus faecalis* to *Leuconostoc oenos*. *FEMS Microbiol. Lett.* **1996**, *135*, 179–185. [CrossRef]
52. Mesas, J.M.; Rodriguez, M.C.; Alegre, M.T. Characterization of lactic acid bacteria from musts and wines of three consecutive vintages of Ribeira Sacra. *Lett. Appl. Microbiol.* **2011**, *52*, 258–268. [CrossRef]
53. Groth, A.C.; Calos, M.P. Phage integrases: Biology and applications. *J. Mol. Biol.* **2004**, *335*, 667–678. [CrossRef]
54. Jaomanjaka, F.; Claisse, O.; Blanche-Barbat, M.; Petrel, M.; Ballestra, P.; Le Marrec, C. Characterization of a new virulent phage infecting the lactic acid bacterium *Oenococcus oeni*. *Food Microbiol.* **2016**, *54*, 167–177. [CrossRef]
55. Auvray, F.; Coddeville, M.; Ordonez, R.C.; Ritzenthaler, P. Unusual structure of the *attB* site of the site-specific recombination system of *Lactobacillus delbrueckii* bacteriophage mv4. *J. Bacteriol.* **1999**, *181*, 7385–7389. [CrossRef]
56. Campbell, A. Phage integration and chromosome structure. A personal history. *Annu. Rev. Genet.* **2007**, *41*, 1–11. [CrossRef] [PubMed]
57. Landy, A.; Ross, W. Viral integration and excision: Structure of the lambda att sites. *Science* **1977**, *197*, 1147–1160. [CrossRef] [PubMed]
58. Wojciak, J.M.; Sarkar, D.; Landy, A.; Clubb, R.T. Arm-site binding by lambda -integrase: Solution structure and functional characterization of its amino-terminal domain. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 3434–3439. [CrossRef]
59. Cho, E.H.; Gumport, R.I.; Gardner, J.F. Interactions between integrase and excisionase in the phage lambda excisive nucleoprotein complex. *J. Bacteriol.* **2002**, *184*, 5200–5203. [CrossRef]
60. Suzuki, S.; Yoshikawa, M.; Imamura, D.; Abe, K.; Eichenberger, P.; Sato, T. Compatibility of Site-Specific Recombination Units between Mobile Genetic Elements. *iScience* **2020**, *23*, 100805. [CrossRef] [PubMed]
61. Rutkai, E.; György, A.; Dorgai, L.; Weisberg, R.A. Role of Secondary Attachment Sites in Changing the Specificity of Site-Specific Recombination. *J. Bacteriol.* **2006**, *188*, 3409–3411. [CrossRef] [PubMed]
62. Nakagawa, I.; Kurokawa, K.; Yamashita, A.; Nakata, M.; Tomiyasu, Y.; Okahashi, N.; Kawabata, S.; Yamazaki, K.; Shiba, T.; Yasunaga, T.; et al. Genome sequence of an M3 strain of *Streptococcus pyogenes* reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution. *Genome Res.* **2003**, *13*, 1042–1055. [CrossRef] [PubMed]
63. Philippe, C.; Chaïb, A.; Jaomanjaka, F.; Claisse, O.; Lucas, P.M.; Samot, J.; Cambillau, C.; Le Marrec, C. Characterization of the First Virulent Phage Infecting *Oenococcus oeni*, the Queen of the Cellars. *Front. Microbiol.* **2021**, *11*, 596541. [CrossRef] [PubMed]