Exopolysaccharide (EPS) synthesis by oenococcus oeni: from genes to phenotypes
Maria Dimopoulou, Marlène Vuillemin, Hugo Campbell-Sills, Patrick Lucas, Patricia Ballestra, Cecile Miot-Sertier, Marion Favier, Joana Coulon, Virginie Moine, Thierry Doco, et al.

To cite this version:
Maria Dimopoulou, Marlène Vuillemin, Hugo Campbell-Sills, Patrick Lucas, Patricia Ballestra, et al.. Exopolysaccharide (EPS) synthesis by oenococcus oeni: from genes to phenotypes. PLoS ONE, Public Library of Science, 2014, 9 (6), 15 p. 10.1371/journal.pone.0098898 . hal-02641554

HAL Id: hal-02641554
https://hal.inrae.fr/hal-02641554
Submitted on 28 May 2020

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.
Exopolysaccharide (EPS) Synthesis by Oenococcus oeni: From Genes to Phenotypes

Maria Dimopoulou1, Marlène Vuillemin2, Hugo Campbell-Sills1, Patrick M. Lucas1, Patricia Ballestra1, Cécile Miot-Sertier1, Marion Favier3, Joana Coulon3, Virginie Moine3, Thierry Doco4, Maryline Roques4, Pascale Williams4, Melina Petrel5, Etienne Gontier5, Claire Mouliès2, Magali Remaud-Simeon2, Marguerite Dols-Lafargue1*

1 Université de Bordeaux, Institut polytechnique de Bordeaux, ISVV, EA 4577, Unité de recherche Oenologie, INRA USC 1366, Villenave d’Ornon, France, 2 Université de Toulouse, INSA, UPS, INRA, CNRS, LISBP, Toulouse, France, 3 BioLaffort, research subsidiary of the Laffort Group, Bordeaux, France, 4 INRA, UMR1083, Sciences pour l’œnologie, Montpellier, France, 5 Université de Bordeaux, Bordeaux Imaging Center, UMS 3420 CNRS - US4 INSERM, Bordeaux, France

Abstract

Oenococcus oeni is the bacterial species which drives malolactic fermentation in wine. The analysis of 50 genomic sequences of O. oeni (14 already available and 36 newly sequenced ones) provided an inventory of the genes potentially involved in exopolysaccharide (EPS) biosynthesis. The loci identified are: two gene clusters named eps1 and eps2, three isolated glycoside-hydrolase genes named dsrO, dsrV and levO, and three isolated glycosyltransferase genes named gtf, it3, it4. The isolated genes were present or absent depending on the strain and the eps gene clusters composition diverged from one strain to another. The soluble and capsular EPS production capacity of several strains was examined after growth in different culture media and the EPS structure was determined. Genotype to phenotype correlations showed that several EPS biosynthetic pathways were active and complementary in O. oeni. Can be distinguished: (i) a Wzy -dependent synthetic pathway, allowing the production of heteropolysaccharides made of glucose, galactose and rhamnose, mainly in a capsular form, (ii) a glucan synthase pathway (Gtf), involved in β-glucan synthesis in a free and a cell-associated form, giving a ropy phenotype to growth media and (iii) homopolysaccharide synthesis from sucrose (α-glucan or β-fructan) by glycoside-hydrolases of the GH70 and GH68 families. The eps gene distribution on the phylogenetic tree was examined. Fifty out of 50 studied genomes possessed several genes dedicated to EPS metabolism. This suggests that these polymers are important for the adaptation of O. oeni to its specific ecological niche, wine and possibly contribute to the technological performance of malolactic starters.


Editor: Benoît Fouligné, Institut Pasteur de Lille, France
Received February 14, 2014; Accepted May 8, 2014; Published June 5, 2014
Copyright: © 2014 Dimopoulou et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
Funding: ANR-10-ALIA-003 (www.agence-nationale-recherche.fr). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
Competing Interests: M. Favier, J. Coulon and V. Moine are employed by BioLaffort. However, this does not alter the authors’ adherence to PLOS ONE policies on sharing data and materials.
* E-mail: dols@enscbp.fr

Introduction

Oenococcus oeni, formerly Leuconostoc oenos is the bacterial species which most frequently drives malolactic fermentation (MLF) in wine. Nowadays, MLF is recommended for most red wines (and sometimes for white ones), especially when they are meant to age in the bottle. The transformation of malic acid into lactic acid. However, many other metabolic transformations occur during MLF which undoubtedly have a major effect on wine quality. In order to better control MLF, the use of O. oeni as a malolactic starter was proposed early [4]. Wines are inoculated with selected O. oeni strains at the end of or after alcoholic fermentation. However, O. oeni strains strongly differ regarding their respective ability to survive and conduct MLF after inoculation in wine [5–6]. Comparative genomic as well as less global studies led to identify genes with potential technological interest [2,7–12]. Among the metabolic equipments which could explain the different tolerance to inoculation in wine, the biosynthesis of exopolysaccharides (EPS) was recently examined through genomic studies [12], in wine [13] or through the functional study of specific glucan-synthase [14]. EPS are extracellular polymers composed of sugar monomers. With the few O. oeni strains studied, the soluble EPS yields and the EPS monomer composition vary depending on the strain and/or on the growth medium composition [15]. Actually, O. oeni is able to synthesize both homo and heteropolysaccharides, via distinct metabolic pathways [16]. Most of the time, the medium viscosity is unaltered after EPS synthesis, with the exception of ropy strains which produce β-glucan [13–14,16–18]. Considering that O. oeni genome has a limited size (<1.8 Mb), whole genome sequencing appeared to be the best strategy to rapidly assess the diversity of genes associated with EPS biosynthesis present in the O. oeni pangenome. We therefore analyzed the 14 genomic sequences available [12], and 36 new sequenced ones. The 50 strains studied displayed divergent EPS production level and represented different genetic groups in the O. oeni species [19–20]. Glycosyltransferase, glycoside-hydrolase and sugar nucleotide precursor biosynthetic genes were identified and
the gene cluster organisation was investigated. The link between *eps* genes and the observed EPS phenotypes as well as the *eps* gene distribution on the *O. oeni* species phylogenetic tree were examined.

**Materials and Methods**

**Strains**

The names of the *O. oeni* strains studied and their origin are presented in Table 1. *Lactococcus lactis* IL1403 was also used for developing the method for capsule observation by electronic microscopy.

**Genome Screening, eps Gene Identification and Nomenclature**

Genomic sequences were recovered from databases or produced by GtT-PlaGe Genotoul (Castanet Tolosan France) and Mackrogen (Seoul Korea) (unpublished). All 36 new sequences were annotated by RAST (Rapid Annotation using Subsystem Technology, rast.nmpdr.org) and Kaas (KEGG Automatic Annotation Server) [21]. These sequences have been deposited at DDBJ/EMBL/GenBank under the accession numbers listed in Table 1. The versions described in this paper for *eps* gene content are versions XXXX01000000.

Multilocus sequence typing (MLST) was performed for all strains according to the procedure described by Bilhère et al. [19] with some modifications. The sequence type (ST) of each strain was constructed from six housekeeping genes: *gyrB*, *gfp*, *pgm*, *dnaE*, *purK* and *rpoB* whose sequences were obtained by genome analysis in Seed Viewer application of RAST. Sequence treatment was performed by using BioEdit 7.2.3 and the phylogenetic tree was constructed by the neighbor-joining method with a Kimura two-parameter distance model, using MEGA 4 software [22]. Bootstrap values were obtained after 1,000 iterations.

From the 3 genomes sequences publicly available at the beginning of our work (genomes of strains *O. oeni* PSU-1, ATCC BAA-1163 and AWRI B429), we created a database of 82 protein sequences (Table S1, panel initial database), potentially associated with the EPS metabolism including glycosyltransferases, flippases (wzx) and polymerases (wzy) but also glycoside-hydrolases and protein sequences involved in the synthesis of precursors (sugar nucleotides). The 47 other annotated genome sequences were then analyzed for the presence of orthologs of these 82 proteins (BLASTP). Once an ortholog was identified, the gene genomic content were obtained by genome analysis in Seed Viewer application of RAST. Sequence treatment was performed by using BioEdit 7.2.3 and the phylogenetic tree was constructed by the neighbor-joining method with a Kimura two-parameter distance model, using MEGA 4 software [22]. Bootstrap values were obtained after 1,000 iterations.

From the 3 genomes sequences publicly available at the beginning of our work (genomes of strains *O. oeni* PSU-1, ATCC BAA-1163 and AWRI B429), we created a database of 82 protein sequences (Table S1, panel initial database), potentially associated with the EPS metabolism including glycosyltransferases, flippases (wzx) and polymerases (wzy) but also glycoside-hydrolases and protein sequences involved in the synthesis of precursors (sugar nucleotides). The 47 other annotated genome sequences were then analyzed for the presence of orthologs of these 82 proteins (BLASTP). Once an ortholog was identified, the gene genomic content were obtained by genome analysis in Seed Viewer application of RAST. Sequence treatment was performed by using BioEdit 7.2.3 and the phylogenetic tree was constructed by the neighbor-joining method with a Kimura two-parameter distance model, using MEGA 4 software [22]. Bootstrap values were obtained after 1,000 iterations.

**Materials and Methods**

**ESYS Synthesis and Quantification**

After a two-week growth in SMD medium at 25°C without agitation, the soluble EPS concentration was measured. The whole culture medium was centrifuged (8,000xg, 5 min, 4°C), and the pellet was removed. Three volumes of ethanol-HCl 1 N (95-5) were added to the supernatant to precipitate the polysaccharides. The tubes were let to stand for 24 hours at 4°C. Then, they were centrifuged (18,000xg, 5 min, 4°C), and the pellet was washed with ethanol (80%vol), centrifuged again, dried for 20 min at 65°C and dissolved in distilled water. The amount of neutral polysaccharides was determined by the anthrone sulfuric acid method [24], using glucose as the standard. For each sample, the polymer precipitation and assays were done in triplicate.

**Immunooagglutination and Capsule Observation**

To visualize the bacterial capsule, 10 μl of cell suspension (one week grape juice or SMD culture broth) were deposited on a microscope slide and mixed with 20% nigrosine aqueous solution and let to dry (5 min). Afterwards, 10 μl of 1% crystal violet solution was added and the slide was examined under Olympus BX51 microscope (×100, under oil immersion). The capsule appeared as a white halo around the cells. The β-glucan layer was not sufficiently compact to be visualized by this method. As a result, agglutination tests were performed using *S. pneumoniae* type 57-specific antiserum, as previously reported [14]. Four microliters of antiserum were spotted on a slide with 20 μl of culture broth and incubated 30 min at 4°C before observation using phase contrast microscopy.

For transmission electron microscopy (TEM), bacteria were fixed for 2 hours in 0.1 M sodium cacodylate buffer (pH 7.2) containing 2% glutaraldehyde, at room temperature. Fixed bacteria were stored at 4°C in the fixative solution. They were rinsed in cacodylate buffer, then in 1% gelatin and postfixed (i) with 1% osmium tetroxide containing 1.5% potassium cyanoferrate and (ii) with 3% uranyl acetate at 4°C. They were gradually dehydrated in ethanol (30% to 100%) and embedded in Epon. Thin sections (70 nm) were collected on 150-mesh copper grids, before examination with a Hitachi H7650 TEM. Negative staining and TEM observation gave the same results (presence or absence of capsule) for all the strains examined.

**EPS Purification and Structural Analysis**

For capsule structure determination, 500 mL of SMD-glucose culture medium was centrifuged and the pellet was washed twice with PBS buffer (NaCl 137 mM, KCl 2.7 mM, Na2HPO4 10 mM, pH 7). Then, the pellet was washed with 100 ml of ultrapure water and the cell walls were recovered by centrifugation (6000xg, 4°C, 20 min) and freeze dried. The capuslar polysaccharide was extracted by the method of Mendes [25] and purified as described by Pichon et al. [26]. The identity of the EPS was confirmed by 1D and 2D NMR spectra using a Bruker AVANCE 500MHz spectrometer and by MALDI-TOF-MS. Aminosugars and sugars were identified by GC-MS after derivatization of the samples with triethylsilyl trifluoroacetamide [27]. The molecular weight of the EPS was measured by MALDI-TOF-MS [28]. The degree of sialylation was calculated as the ratio of the area of the peak corresponding to the monosialylated form to the area of the peak corresponding to the non-sialylated form. The peak molecular weights of the EPS were calculated by using the Molecular Operating Environment (MOE). The E. coli K1 capsule was also used as a control.

**EPS Purification and Structural Analysis**

For capsule structure determination, 500 mL of SMD-glucose culture medium was centrifuged and the pellet was washed twice with PBS buffer (NaCl 137 mM, KCl 2.7 mM, Na2HPO4 10 mM, pH 7). Then, the pellet was washed with 100 ml of ultrapure water and the cell walls were recovered by centrifugation (6000xg, 4°C, 20 min) and freeze dried. The capuslar polysaccharide was extracted by the method of Mendes [25] and purified as described by Pichon et al. [26]. The identity of the EPS was confirmed by 1D and 2D NMR spectra using a Bruker AVANCE 500MHz spectrometer and by MALDI-TOF-MS. Aminosugars and sugars were identified by GC-MS after derivatization of the samples with triethylsilyl trifluoroacetamide [27]. The molecular weight of the EPS was measured by MALDI-TOF-MS [28]. The degree of sialylation was calculated as the ratio of the area of the peak corresponding to the monosialylated form to the area of the peak corresponding to the non-sialylated form. The peak molecular weights of the EPS were calculated by using the Molecular Operating Environment (MOE). The E. coli K1 capsule was also used as a control.

**Exopolysaccharide Synthesis by O. oeni**

*O. oeni* was propagated either in Grape juice medium [15] or in a semi defined (SMD) medium specifically developed for EPS production by *O. oeni*. The SMD medium contained: (base) casamino acids 10 g/L, sodium acetate 3.4 g/L, KH2PO4 1 g/L, MgSO4·7 H2O 0.1 g/L, MnSO4·H2O 4 H2O 0.1 g/L, ammonium citrate 2.7 g/L, lactotryptone 3 g/L, malate 3 g/L, yeast nitrogen base 6.7 g/L, adenine, uracil, thymine, guanine 5 mg/L each, and a carbohydrate (either glucose 20 g/L or glucose and sucrose, 10 g/L each). The pH was adjusted to 3.0. The carbohydrate solutions were prepared as 10X solutions and were sterilized 20 min at 121°C, while the base was prepared as a 2X solution and sterilized by filtration (0.2 μm cut off). *L. lactis* was propagated in MRS medium [15].
Table 1. List and origin of the strains studied.

<table>
<thead>
<tr>
<th>Strain name&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Collection&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Origin/commercial name</th>
<th>Accession number&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>0205</td>
<td>IOEB</td>
<td>Champagne isolate</td>
<td>AZHH00000000</td>
</tr>
<tr>
<td>0501</td>
<td>IOEB</td>
<td>Red wine France</td>
<td>AZIP00000000</td>
</tr>
<tr>
<td>0502</td>
<td>IOEB</td>
<td>French isolate</td>
<td>AZKL00000000</td>
</tr>
<tr>
<td>0607</td>
<td>IOEB</td>
<td>French isolate</td>
<td>AZKK00000000</td>
</tr>
<tr>
<td>0608</td>
<td>IOEB</td>
<td>French isolate</td>
<td>AZKJ00000000</td>
</tr>
<tr>
<td>1491</td>
<td>IOEB</td>
<td>Red wine France</td>
<td>AZLG00000000</td>
</tr>
<tr>
<td>277</td>
<td>IOEB-S commercial</td>
<td>SB3, Laffort France</td>
<td>AZKD00000000</td>
</tr>
<tr>
<td>436a</td>
<td>IOEB-S</td>
<td>Red wine Bordeaux France</td>
<td>AZLS00000000</td>
</tr>
<tr>
<td>540</td>
<td>IOEB-S commercial</td>
<td>450 PreAc, Laffort, France</td>
<td>AZLD00000000</td>
</tr>
<tr>
<td>8417</td>
<td>IOEB</td>
<td>Ropy red wine</td>
<td>AZKX00000000</td>
</tr>
<tr>
<td>9304</td>
<td>IOEB</td>
<td>Cider France</td>
<td>AZKX00000000</td>
</tr>
<tr>
<td>9517</td>
<td>IOEB</td>
<td>Floc de Gascogne France</td>
<td>AZKG00000000</td>
</tr>
<tr>
<td>9803</td>
<td>IOEB</td>
<td>Red wine France</td>
<td>AZKF00000000</td>
</tr>
<tr>
<td>9805</td>
<td>IOEB</td>
<td>Red wine France</td>
<td>AZKE00000000</td>
</tr>
<tr>
<td>ATCC BAA-1163</td>
<td>ATCC</td>
<td>Red wine, France</td>
<td>AAUV00000000*</td>
</tr>
<tr>
<td>B10</td>
<td>IOEB</td>
<td>French isolate</td>
<td>AZJW00000000</td>
</tr>
<tr>
<td>B129*</td>
<td>AWRI</td>
<td>DSM 20252/ATCC23279, Red wine France</td>
<td>AJPT00000000*</td>
</tr>
<tr>
<td>B16</td>
<td>IOEB commercial</td>
<td>B16, Laffort France</td>
<td>AZKC00000000</td>
</tr>
<tr>
<td>B202*</td>
<td>AWRI</td>
<td>Australian isolate</td>
<td>AJTO00000000*</td>
</tr>
<tr>
<td>B304*</td>
<td>AWRI</td>
<td>Australian isolate</td>
<td>AUIJ00000000*</td>
</tr>
<tr>
<td>B318*</td>
<td>AWRI</td>
<td>NCDO 1884, Australia</td>
<td>ALAD00000000*</td>
</tr>
<tr>
<td>B418</td>
<td>AWRI</td>
<td>MCW Lallemand</td>
<td>ALAE00000000*</td>
</tr>
<tr>
<td>B419</td>
<td>AWRI</td>
<td>Lalvin EQ54 Lallemand</td>
<td>AALF00000000*</td>
</tr>
<tr>
<td>B422</td>
<td>AWRI</td>
<td>Vinflora CHR35, Chr. Hansen</td>
<td>ALEG00000000*</td>
</tr>
<tr>
<td>B429</td>
<td>AWRI</td>
<td>Lalvin VP41 Lallemand</td>
<td>ACSE00000000*</td>
</tr>
<tr>
<td>B548</td>
<td>AWRI</td>
<td>BL-01 Lallemand</td>
<td>ALAH00000000*</td>
</tr>
<tr>
<td>B553*</td>
<td>AWRI</td>
<td>Elios-1 Lallemand</td>
<td>ALAI00000000*</td>
</tr>
<tr>
<td>B568*</td>
<td>AWRI</td>
<td>Australian isolate</td>
<td>ALAJ00000000*</td>
</tr>
<tr>
<td>B576*</td>
<td>AWRI</td>
<td>Australian isolate</td>
<td>ALAK00000000*</td>
</tr>
<tr>
<td>C23</td>
<td>IOEB</td>
<td>Cider Normandy France</td>
<td>AAZU00000000</td>
</tr>
<tr>
<td>C28</td>
<td>IOEB</td>
<td>Cider, Bretagne France</td>
<td>AZLE00000000</td>
</tr>
<tr>
<td>C52</td>
<td>IOEB</td>
<td>Cider Normandy France</td>
<td>AZLF00000000</td>
</tr>
<tr>
<td>CiNe</td>
<td>IOEB</td>
<td>Starter CHR Hansen</td>
<td>AZJV00000000</td>
</tr>
<tr>
<td>L18_3</td>
<td>IOEB</td>
<td>Red wine Lebanon</td>
<td>AZLO00000000</td>
</tr>
<tr>
<td>L26_1</td>
<td>IOEB</td>
<td>Lebanese isolate</td>
<td>AZLP00000000</td>
</tr>
<tr>
<td>L40_4</td>
<td>IOEB</td>
<td>Red wine Lebanon</td>
<td>AZLF00000000</td>
</tr>
<tr>
<td>L65_2</td>
<td>IOEB</td>
<td>Red wine, Lebanon</td>
<td>AZLR00000000</td>
</tr>
<tr>
<td>PSU-1 commercial</td>
<td>commercial</td>
<td>Red wine USA</td>
<td>NC_008528*</td>
</tr>
<tr>
<td>S11</td>
<td>S</td>
<td>Sparkling white wine France</td>
<td>AZJX00000000</td>
</tr>
<tr>
<td>S12</td>
<td>S</td>
<td>White wine France</td>
<td>AZLH00000000</td>
</tr>
<tr>
<td>S13</td>
<td>S</td>
<td>Red wine France</td>
<td>AZKB00000000</td>
</tr>
<tr>
<td>S14</td>
<td>S</td>
<td>Red wine France</td>
<td>AZLI00000000</td>
</tr>
<tr>
<td>S15</td>
<td>S</td>
<td>Red wine France</td>
<td>AZLJ00000000</td>
</tr>
<tr>
<td>S161</td>
<td>S, commercial</td>
<td>350 PreAc, Laffort France</td>
<td>AZLM00000000</td>
</tr>
<tr>
<td>S19</td>
<td>S</td>
<td>Red wine France</td>
<td>AZLK00000000</td>
</tr>
<tr>
<td>S22</td>
<td>S</td>
<td>Sparkling white wine Bourgogne France</td>
<td>AZKA00000000</td>
</tr>
<tr>
<td>S23</td>
<td>S</td>
<td>white wine, England</td>
<td>AZLL00000000</td>
</tr>
<tr>
<td>S25</td>
<td>S</td>
<td>Red wine France</td>
<td>AZLZ00000000</td>
</tr>
</tbody>
</table>
The molecular weight distribution of an aqueous solution of freeze dried soluble EPS was established by high-performance size-exclusion chromatography (HPSEC) using a system composed of a 234-Gibson sampling injector (Roissy, France) and an LC-10 AS Shimadzu pump (Kyoto, Japan). HPSEC elution was performed on two serial Shodex OHPAK KB-803 and KB-805 columns (0.8×30 cm; Showa Denko, Japan), connected to an ERC-7512 refractometer (Erma, Japan), at a 1 mL/min flow rate in 0.1M LiNO₃. The apparent molecular weights were calculated from the calibration curve established with a Pullulan calibration kit (Showa Denko, Japan).

Neutral monosaccharides were released after polysaccharides hydrolysis by treatment with 2 M trifluoroacetic acid (120°C, 75 min) [26]. The released monosaccharides were methanlated using methyl sulfinyl carbanion and methyl iodide [27], and converted to their corresponding alditol acetates by treatment with NaBD₄ and then acetylated [28]. The methylated residues were quantified by gas chromatography (GC), using a fused silica DB-225 (210°C) capillary column (30 m × 0.32 mm internal diameter, 0.25 μm film), with hydrogen as the carrier gas, on a Shimadzu GC-2010 gas chromatograph. The alditol acetates were identified from their retention times, by comparison with standards. Neutral sugars amounts were calculated relative to the internal standard (myo-inositol).

The neutral, acidic and amino sugar composition of the EPS was determined after N-acyetylasefication after solvolysis with anhydrous MeOH containing 0.5 M HCl (30°C, 16 h), and gas chromatography of the per-O-trimethylsilylated methyl glycoside derivatives (TMS). The TMS derivatives were separated on two DB-1 capillary columns (30 m × 0.25 mm i.d., 0.25 μm film) (temperature program 120 to 200°C, 1.5°C/min), coupled with a single injector inlet, through a two-holed ferrule, with H₂ as the carrier gas, on a Shimadzu GCMS-QP2010SE gas chromatograph. The outlet of one column was directly connected to a FID (250°C). The second column was connected to a mass detector, via a deactivated fused-silica column (0.25 m × 0.11 μm i.d.). Samples were injected in pulsed split mode, with a 20:1 split ratio. The transfer line to the mass was set at 200°C. Electron Ionization (EI) mass spectra were obtained from m/z 50 to 400 every 0.2 s, in total ion-monitoring mode (200°C ion source temperature, a 60 μA filament emission current and a 70 eV ionization voltage).

The EPS produced on SMD-Glucose-sucrose were also analyzed for glycosidic linkage. Five mg of EPS in 0.5 ml dimethylsulfoxide were methyalted as described above and then hydrolyzed with 2 M trifluoroacetic acid (120°C, 1.15 h). The released methylated monosaccharides were converted to their corresponding alditol acetates. The partially methylated alditol acetates were analyzed by GC-EI-MS on a Shimadzu GCMS-QP2010SE gas chromatograph using a DB-1 capillary column (30 m × 0.25 mm i.d., 0.25 μm film) and the following temperature program: 135°C for 10 min, and rise to 180°C at 1.2°C/min. The transfer line to the mass was set at 280°C. EI mass spectra were obtained from m/z 50 to 400 every 0.2 s, in total ion-monitoring mode (200°C ion source temperature, a 60 μA filament emission current and a 70 eV ionization voltage).

**Results**

**eps Gene Inventory**

**Global analysis.** Many genes potentially associated with EPS biosynthesis were identified: these included glycosyltransferase and glycoside hydrolase genes, either isolated or clustered, and genes associated with the synthesis of nucleotide-sugars or other precursors. These genes are listed in Table S1. Only some of these genes, because (i) their link with EPS metabolism is plausible and (ii) they are not strictly conserved in all the genomes studied, will be presented in detail in this article. All the genes studied were chromosomal (Figure 1). There were two complex heteropolysaccharide clusters, eps1 and eps2, displaying a high density of coding sequences and related to the eps clusters previously described by Dimopoulou et al. [16], genes of glycoside-hydrolases (dsrO, dsrV and levO) and 3 isolated glycosyltransferase genes (gtf, it3 and it4).

All the genes and clusters studied, when present, were always located at the same site on the bacterial chromosome, except the gtf gene which could be found in two different positions in the

---

**Table 1.** Cont.  

<table>
<thead>
<tr>
<th>Strain namea</th>
<th>Collectionb</th>
<th>Origin/commercial name</th>
<th>Accession numberc</th>
</tr>
</thead>
<tbody>
<tr>
<td>528</td>
<td>Commercial</td>
<td>B28 PreAc, Laffort, France</td>
<td>AZJY00000000</td>
</tr>
<tr>
<td>VF</td>
<td>Commercial</td>
<td>Starter VF, Martin Vialatte</td>
<td>AZLM00000000</td>
</tr>
</tbody>
</table>

---

*a the * indicates that the genome sequence was already available in the databases.  
*TCC: American type culture collection; AWR: Australian wine research institute; IOEB: Institut d’Oenologie de Bordeaux, France; S. Sarco, Biolaffort, France.

*The* indicates that the sequence was already available in the databases.

---

*doi:10.1371/journal.pone.0098898.t001*
chromosome (Figure 1). The analysis also indicated that each of the 50 genomes studied was equipped with several distinct genes encoding distinct EPS biosynthetic pathways. This point will be detailed below, locus by locus.

Cluster eps1. All the genomes studied displayed a eps1 cluster. The analysis the 50 eps1 sequences indicated the existence of three related models named A, B and C (Figure 2). Fourteen out of 50 genomes displayed a model A of cluster eps1, 28/50 genomes displayed a model B, and the remaining eight genomes had a model C. When two genomes displayed the same model of eps1, the cluster gene sequences were over 97% conserved.

The three models of cluster eps1 differed by the presence of additional genes and by gene synteny. However, more than half of the genes in the cluster were highly conserved (Figure 2, Table 2). The genes encoding UDP-glucose dehydrogenase (ugd) and galactopyranosidase (glf) were the most conserved ones. The model A was that previously described for strains PSU-1 and BAA-1163 [16]. This was the least complex model of cluster eps1 regarding the glycosyltransferase gene composition (5 genes, Table 2). Model B differed from model A by the presence of five additional genes (woaF, G, H, I and J). Model B therefore encoded seven putative glycosyltransferases, a putative phosphoglyceroltransferase WoaF and a protein with unknown function, WoaM. Moreover, WoaD and WoaE were relatively divergent between models B and A (Table 2). In model C, the gene woaF was present, as in model B, but genes woaC, D, E, G, H, I and J were absent and new genes were present (woaK, L, M and N, Figure 2).

The protein Wzy encoded in model C was highly divergent compared to versions A and B (Table 2).

Whatever the model, the cluster apparently brought all the information necessary for the establishment of a heteropolysaccharide biosynthetic pathway: a priming glycosyltransferase gene woaA, genes encoding glycosyltransferases potentially associated with the synthesis of the repeating unit (woaB to woaN) or to precursor synthesis, glf and ugd. The functional annotation of Ugd, Glf and WoaF suggests the presence of glucuronic acid, phosphoglycerol and galactose in the synthesized product. The wzz gene encoded a protein which exhibited little homology in the data bases, but may participate in the regulation of the biosynthetic pathway (chain length regulation). The cluster also comprised a flippase gene, wzx, and a potent polymerase gene, wzy. Indeed, whatever the model of cluster eps1 considered, the gene wzy was very singular. It may encode a polysaccharide polymerase (Wzy) and, in this case, the cluster encodes a complete heteropolysaccharide biosynthetic pathway. However, the analysis of conserved domains (PFAM hidden Markov models (HMM) Table S1, panel eps1) and the analysis of membrane spanning domains (not shown) suggest that it might rather be a O- antigen ligase (Wzy-C superfamily, WaaL). Enzymes of this family catalyze the binding of polysaccharides moieties of lipopolysaccharide on the oligosaccharide core anchored in the lipid membrane in Gram negative bacteria [29] However, such an activity has never been described in Gram-positive bacteria.

Cluster eps2. Forty-three out of fifty genomes displayed a second heteropolysaccharide cluster eps2. Fifteen models of cluster
encoded a priming glycosyltransferase related to WobA B429 and eps2 were found among the 13 models of cluster gene (Figure 3). The protein WobA B429 displayed 39% identity genomes, the cluster also in that of strains B422, B548, B16 and 0205. In the last four wobA was absent in the genome of strain ATCC BAA-1163, but with WobA S12 and 65% identity WobA PSU1, while forms eps2 related to WobA PSU1 (protein identity 99%), three models were systematically inserted between genes amiO and recP (Figure 3). The end of the gene, most of the models of eps2 cluster displayed the genes encoding the glycosyltransferases potentially involved in the repeating unit synthesis. The polymerase and flippase genes but also genes encoding enzymes involved in precursor synthesis or modification complete the cluster. The 5’ end of this part of the cluster (beyond wobA) was sometimes conserved between genomes (black arrows), whereas the 3’ end was highly divergent (light gray arrows in Figure 3). Indeed, in that 3’ end “gray” zone of cluster eps2, no nucleotide identity was found between models taken in pairs, except for a few flippase genes (wzx, see below). However, function homologies (same PFAM) between encoded proteins were common. The proteins deduced from genes in this 3’-end of the eps2 clusters displayed homologies (35 to 85%) with proteins sequenced from very diverse bacteria: Lactobacillus rhamnosus, Lb casei, Lb fermentum, Lb amylovorus, Lb paracasei, Lb delbrueckii, Lb plantarum, Lb rhamnsis, Streptococcus thermophilus, S. pneumoniae, S. sanguis, S. sanguinis, S. agalactiae, Leuconostoc citreum, Lm mesenteroides, L. lactis, Pediococcus acidilactici, Enterococcus faecalis, Bifidobacterium bifidum, Bacillus coagulans or Bacteroides dorei. Few of these species are encountered in wine environment, but very few wine bacteria genomes have been sequenced and published at the time of this study.

Sequence analysis of the protein sequences deduced from the 15 models of cluster eps2 led to identify (Figure 3, Table S1, panel eps2):

- 3 highly conserved regulatory proteins (AraC, Wzd, Wze),
- 13 distinct polymerase (Wzy), displaying low identity with the sequences in the database. WodC encoded in model 9304 of eps2 may be a 14th polymerase,
- 9 flippases families: B422/PSU1 (99% identity), BAA-1163/9805 (80% identity), 0502/9304/0607/C52/C23 (more than 75% identity), 0501, B429, 9517, S13, 277, S12,
- 3 alleles of priming glycosyltransferases WobA (WobA<sub>PSU1</sub>, WobA<sub>B429</sub>, WobA<sub>S12</sub>),
Figure 3. Comparison of the eps2 gene clusters. In front of each model of cluster eps2, the name of the model strain and the size of the cluster are indicated. When present, the eps2 cluster is always located between recP and amiO (core genome genes in O. oeni chromosome). It displays, in its 5’ end, the three genes araC, wzd and wze, the initial transferase gene wobA (3 different versions), and then, genes specific to each model. The arrows filling indicate the putative function of the encoded proteins. The black and dark gray fillings indicate genes shared by several models of eps2. On the other hand, light gray arrows indicate genes specific to a single model. Groups of strains bearing the same eps2 cluster: Model PSU-1: B418, Model 0502: B10, Model 0607: L26_1, S22, S25, Model B553: L65_2, 9517, Model 277: S15, S161, L18_3, 450, S14, Model 9805: 9803, 8417, Model 9304: C28, Model B429: B202, B304 B318, B568, B576, 0608, C1Ne, S11, S23, S28, model B422: B548, 0205, B16. No eps2: VF, S19, 1491, B129, L40_4, 436a, B419. doi:10.1371/journal.pone.0098898.g003

Exopolysaccharide Synthesis by O. oeni

PLOS ONE | www.plosone.org 7 June 2014 | Volume 9 | Issue 6 | e98898
5 putative rhamnosyltransferases WobB, WobF, WobJ, WobS and WobU (1 GT-1, 4 GT-2),
4 putative galactosyltransferases, WocK, WocS, WodQ and WodS (1 GT-1, 2 GT-2, 1 GT-28),
3 putative choline phosphotransferases (LicD_{277}, LicD_{223}, LicD_{352}),
1 putative glucosyltransferase, WobE,
53 glycosyltransferases, whose substrate specificity could not be predicted by sequence analysis and, among them, 24 glycosyltransferases classified in GT-2, 19 in GT-1, 2 in GT-4 and 8 not associated with a CAZy family,
4 putative acetyltransferases and 2 putative pyruvyltransferases,
3 UDP-glucose-dehydrogenase (UgdB_{355}, Ugd_{401}, Ugd_{512}), 2 glycerol-3-P-cytidyltransferase (TagD_{277}, TagD_{322}), 1 nucleotidyl transferase (Abp_{232}) and 1 epimerase (Abp_{232}),

The substrate specificity prediction for glycosyltransferases and others enzymes encoded in clusters eps1 and eps2 suggests that the monomers found in the heteropolysaccharides produced by O. oeni may be different from one strain to the other. These heteropolysaccharides may be made of either galactose, rhamnose, glucose and/or glucuronic acid. Furthermore, they may be substituted by acetate, pyruvate, choline and glycerol. Other monomers may also be present, given the high proportion of glycosyltransferases whose protein sequence did not enable to predict their substrate specificity. Nevertheless, the strong similarity between the flipases encoded by different models of cluster eps2 suggests that the repeating units transported may be of relatively close composition or structure, unless these flipases are sufficiently flexible to transport different oligosaccharide structures.

Precursors. Beyond the substrate specificity of the glycosyltransferases in the eps clusters, the precursors biosynthetic pathways may also limit the variety of monomers encountered in O. oeni heteropolysaccharides [30–31]. It is generally accepted that the monomers are transferred from sugar nucleotides (NDP-linked), except for acetyl and pyruvyl which are respectively transferred from acetyl-CoA and phosphoenolpyruvate (PEP). The genes associated with the biosynthesis of these different precursors have been sought in the different genomes (Table S1, panel precursors). Most of these genes were located outside the eps1 and 2 clusters and formed part of the core genome. Thus, as indicated in Figure 4, all the strains studied were equipped to synthesize PEP, acetyl-CoA, UDP-glucose, UDP-galactopyranose and UDP-galactofuranose, dTDP-rhamnose and dTDP-glucose, UDP-glucuronate and, provided that phosphoglucomutase is able to catalyze the conversion of glucosamine-6-phosphate to glucosamine-1-phosphate, UDP-N-acetylglucosamine and UDP-N-acetylgalactosamine.

On the other hand, only a few strains were apparently able to produce CDP-glycerol (protein TagD provided by eps2 model C52) or UDP-N-Acetyl mannosamine (Mna provided by eps2 model C52). Regarding the biosynthesis of NDP-arabinitol, the genes abp1 and abp2 were found in the C52 genome (in cluster eps2) but the deduced proteins exhibited moderate identities with proteins Abp1 and Abp2 found in the databases (37% and 30%). Finally, the biosynthetic pathway for CDP-choline (LicA and LicC) was not found in any of the studied genomes, although three models of cluster eps2 (5 strains involved) encoded a choline phosphotransferase (LicD). Nevertheless, we cannot exclude that these functions are performed by highly divergent proteins in O. oeni.

Additional glycosyltransferase genes. Another element may contribute to the modulation of the structure of the EPS produced by O. oeni: the presence of additional glycosyltransferase genes, outside eps1 and eps2 clusters. However, most of the additional glycosyltransferase genes studied formed part of the core genome (Table S1, panel additional glycosyltransferases). It should be noted, among these highly conserved glycosyltransferase genes, the presence of a priming glycosyltransferase gene (gtf) that could complement truncated eps clusters such as the BAA-1163 eps2 model.

Other genes were present in a smaller number of genomes. Thus, another putative gene of priming glycosyltransferase (gtf) was present in 8/50 genomes. The analysis of adjacent genes indicated that the acquisition of this gene was probably related to a phage attack (gene in a phage remnant). Furthermore, 5 out of 50 genomes encoded a processive glucosyltransferase, Gif, 97% identical to the glucosyltransferase described in Pedicoccus parvulus IOEB 0801, for the biosynthesis of β-1,3-β-1,2 glucan associated with wine ripeness [17,32]. The gif gene of O. oeni IOEB 0203 was previously characterized [14] but its exact location on the chromosome and its presence in the 4 other genomes were discovered in the present study. Two separate insertion sites were identified for gif (Figure 1). The gene is located within a 15.5 kb insert (phage remnant) in the genome of strains B422, B548, 0205 and B16. In 0502 genome, the gif gene was inserted in a potentially mobile prophage (40.9 kb insert).

Glycoside-hydrolases. Three glycoside hydrolases genes were identified. The first one, dsrO, was present in 49 genomes and always inserted in the same site on the chromosome (Figure 1). The entire sequence of this gene extended to 4420 nt (Figure 5). Point mutations could however shorten it, and modify the activity of the proteins produced. For example, for 10 out of 50 strains, dsrO had a stop codon at position 3305 nt, still generating a potentially active protein –as codons for amino acids of the catalytic triad were conserved [33–34]. For 4 strains out of 50, two stop codons in the sequence produced three ORFs, probably encoding inactive DsrO protein fragments. The protein DsrO was 50% more than 90% conserved in the area preceding the mutaiton. In its long form (1475aa), it displayed 72% identity with the dextranucrase DsrP produced by Leuconostoc mesenteroides IBT-PQ (NCBI AAS79426.1) [35].

Eleven out of 50 genomes displayed an additional dextranucrase pseudogene (dsrV), whose sequence was 90% identical (100% coverage) between the genomes displaying it. However, the deduced protein was always truncated in the catalytic site, and may therefore be inactive in all cases (Figure 5). The position of the truncation varied depending on the strain studied. The identity between the genes dsrO and dsrV was 50%.

Thirteen out of 50 genomes had a levansucrase gene (levO), whose sequence was 98% identical between the strains displaying it. In strains 9304, C28 and S13, levO was cut prematurely, and most likely encoded an inactive enzyme. LevO displayed 49% identity with the putative levansucrase identified in Oenococcus kitaharae DSM177330 (WP_007744218.1), and 36% identity with the levansucrase LevS, produced by Leuconostoc mesenteroides B-512 F, characterized in 2006 [36].

Although present in a small number of genomes, and levO and dsrV genes were always inserted at the same site on the chromosome (Figure 1). Analysis of adjacent genes indicated the acquisition of dsrV could be linked to a phage attack (remnant) and rearrangements due to transposases. Regarding levO, no trace of
Figure 4. Putative precursor biosynthetic pathways active in *O. oeni* deduced from genome analysis. The enzyme full names and the accession numbers of reference proteins are shown in Table S1 (panel precursors). The solid arrows indicate the central pathways (glucose 6-P to xylulose-5-P and PEP and acetyl-CoA) and the pathways potentially active in all the strains studied, as the associated enzymes are encoded by the 50 genomes studied. The dashed arrows indicate pathways putatively active in a smaller number of strains. The EPS monomer precursors potentially available in all the strains studied are boxed in solid lines, while the precursors putatively available in a limited number of strains are boxed with dotted lines. "?" indicate metabolic steps for which no enzyme was identified from the genome analyses. P: phosphate, CoA : coenzyme-A, NDP : nucleotidyl-diphosphate, CDP : cytidyl-diphosphate, UDP : uridine-diphosphate; GDP : guanosine-diphosphate, dTDP : deoxythymidine diphasphate, Glc : glucose, Fru : fructose, GlcA : glucuronic acid, Gal : galactose, Galp : galactopyranose, Gal : galactofuranose, LicA : choline kinase, LicC : choline cytidylyltransferase Lrha, L-rhamnose, GIN : glucosamine, N-Ac-Glc : N-acetyl glucosamine, N-Ac-Gal : N-acetyl-galactosamine, N-Ac-Man : N-acetyl-mannosamine, G-A-P : glyceraldehyde 3-phosphate, DHAP: dihydroxyacetone phosphate, PEP : phosphoenolpyruvate.

doi:10.1371/journal.pone.0098898.g004

Figure 5. Genetic organization of *O. oeni* chromosome regions harboring *dsrO* and *dsrV* genes. Example of strains *O. oeni* PSU-1, BAA-1163, 0607 and 277. The strain 277 also displays a *dsrO* gene, similar to that found in *O. oeni* PSU-1.

doi:10.1371/journal.pone.0098898.g005
mobile element nearby could explain the mode of acquisition of the gene.

Distribution of eps Genes and Phylogenetic tree

The 50 genome sequences were used for MLST typing using 6 housekeeping genes in order to construct a consensus dendrogram. The strains distributed into two main phylogenroups (A and B), as previously described [11,19–20]. The repartition of the eps genes and EPS phenotype on this dendrogram was then examined (Figure 6). All genomes in the branch B, except C52, displayed a model A of cluster eps1, while genomes in the branch A displayed the three models of cluster eps1 (A, B or C). The strains having levO or the same version of dsrO were grouped on the phylogenetic tree. In contrast, the strains carrying gft, dsrV or it4, putatively acquired via phage attack, were not grouped.

Regarding cluster eps2, strains that carried the same eps2 model were generally grouped on the tree. For example, the 11 strains having a B429 model were all on the same branch. In other cases, strains with the same eps2 are far apart on the tree: for example, strains displaying model PSU-1 or 0502 of eps2 could belong to the A or B branches of the tree. In addition, strains belonging to remote subdivisions in branch A displayed the model 277 of eps2 (450, S14, S161, L18_3, S15 and 277). In these cases, the acquisition of the eps2 cluster may result from distinct events in the strains considered.

Some links between the eps loci appeared on the dendrogram. Actually, although strains with eps2 model 277 or model 0501 sometimes have a model A of cluster eps1 (450 or 0501), sometimes a model B of cluster eps1 (277, S15, S161, L18_3 and B10), most of the time, when two genomes displayed the same cluster eps2, they also had the same eps1. Indeed, all the genomes with a cluster eps2 model B429 or 0607 displayed a model B of cluster eps1, and all the genomes with a cluster eps2 model 9805 or PSU-1 displayed a model A of cluster eps1, even if they are far apart on the phylogenetic tree. Furthermore, genomes with model C of cluster eps1 systematically had a truncated or absent cluster eps2. In addition, genomes B422, B548, B16 and 0205, in which eps2 cluster was strongly truncated (5.4 kb), were also those whose gft gene was located in a phage remnant. The four strains, all from Champagne region [20], were grouped on the dendrogram. They may have diverged after the acquisition of their eps genes. In addition, these 4 genomes, gft may be “stabilized” compared to the genome 0502 which displayed gft in a prophage and also a non truncated eps2 cluster.

Links between eps Genes and EPS Phenotypes

O. oeni is not amenable to genetic transformation. The consequence is that evidence for phenotype cannot be obtained by gene inactivation. As a result, we analyzed the phenotype of a high number of strains, in order to identify potent links with the identified genotypes. Previous work suggested that, during growth in the presence of glucose as the sole carbon substrate, the EPS synthetic routes using nucleotide sugars were the sole active (Wzy dependent pathway and Gtf synthase pathway), whereas, in the presence of sucrose, the action of glycoside-hydrolases supplement the bacterial biosynthetic capabilities [16]. Phenotypes were therefore studied in the presence of glucose alone or in the presence of glucose and sucrose, most of the O. oeni strains studied being unable to use sucrose as a growth substrate [37–38].

In glucose-only medium, the strains studied produced low amounts of soluble EPS (<80 mg/l) with the exception of strains S15, 277 and of the 5 strains carrying the gft gene (B422, B548, B16, 0205, and 0502), for which the medium also became ropy (Figure 6). The strain IOEB2005 is already known to produce β-glucan [14]. The 4 other ropy strains agglutinated in the presence of antibody targeting the β-glucan (not shown) indicating that they also produced this specific polymer. Except for these ropy strains, it was difficult to establish a link between the concentration of soluble EPS observed after growth in SMD-Glucose and the eps gene variants (Figure 6).

The monomer composition of the few soluble EPS produced on SMD-Glucose was investigated for a selection of 10 strains. All the genomes of the strains studied displayed eps1 and eps2 clusters. The strains 9803, 9805, PSU-1 9304 and S13 displayed a model A of eps1, while the others strains examined displayed a model B. Regarding eps2, the strains S11 and B429 had the same genotype (model B-429), the strains 9803 and 9805 had the same genotype (model 9805), and the others ones (9304–model 9304, S13–model S13,- S22-model 0607–, PSU-1-model PSU-1–, 9517-model B533- and 277–model 277-) displayed different genotypes (figure 6). Soluble polysaccharides obtained after growth in SMD-glucose medium were of moderate size (less than 400 kDa). Whatever the strain studied, the soluble EPS produced on SMD-glucose medium only contained glucose, galactose and rhamnose. No trace of osamine, pyruvate, acetate, glyceral or uronic acid was detected.

The low level of EPS production on SMD-glucose prompted us to look for the presence of capsular polysaccharides. Indeed, after growth on either SMD-glucose or grape juice medium, most of the studied bacteria appeared encapsulated (Figure 6). Only the bacteria having a highly truncated or no eps2 cluster showed no capsule, whatever the model of cluster eps2 they displayed: model B (1491 or L40,4) or model C (B129, 436a, B419, VF, B422, B16, B548 or 0205). Observed by transmission electron microscopy, this capsule was thicker or thinner depending on the strain (Figure 7). Monomer composition analysis of the capsular EPS of strains 9304, S28 and S11 gave the following results: 9304 (Galectose : Glucose : Rhamnose, 68.4: 15.2: 6.9), S28 (Galactose : Glucose : Rhamnose, 41.7: 35.2: 11.1) and S11 (Galactose : Glucose : Rhamnose, 41.2: 31.2: 20.7). The strains S28 and S11, which displayed the same eps1 genotype, produced capsular polysaccharides with close monomer composition compared to strain 9304 which displayed a different eps1 genotype.

The addition of sucrose to the medium induced a marked overproduction of exopolysaccharides with some strains (Figure 6), although 75% did not use sucrose as a growth substrate. The EPS
produced in the presence of sucrose being considerably more abundant, more precise structure analyses could be made (Table 3). First, analysis of the culture supernatants by size exclusion chromatography indicated that the addition of sucrose to the culture medium induced the appearance of a peak corresponding to additional polymers of very high molecular weight (6,000 to 100,000 kDa), with all the strains examined, except strain S25. This last strain was the only one in Table 3 which did not encode a functional glycoside-hydrolase. The structure of the high molecular weight polymer was determined. In all cases, the peak contained a homopolysaccharide or a homopolysaccharide mixture. All strains having a functional dsrO gene (gene length ≥...
3303 nt) produced a 1,6 linked glucan displaying about 5% 1,3 branches. Hydrolysis of the polymer by dextranase confirmed this was an α-glucan (dextran). Besides dextran, strains BAA-1163 and 0501 produced a 2,6-bound fructan. This fructan contained links with β configuration (Vuillemin, unpublished data).

The strains which were not able to produce EPS from sucrose displayed different glycoside-hydrolase genotype and links between genotype and phenotype were not obvious. Indeed, the lack of EPS synthesis from sucrose is coherent in the case of strains with only a truncated dextranucrase *dsrO* (strains 0607, S22, S25, L26_1). However, it cannot be explained, for many others, by the absence or mutation of glycoside-hydrolase genes (i.e. in some strains with a *dsrO* gene 3303 to 4420 nt long, such as GiNe, 0608, S14 and many others, Figure 6).

### Discussion

*Oenococcus oeni*, which drives malolactic fermentation in most wines (especially red ones) and ciders, is very rarely encountered elsewhere or at other stages of winemaking. This is a unique and perfectly specialized bacteria [9]. The analysis of 50 genomes of *O. oeni* shows that genes dedicated to EPS metabolism are distributed all around the chromosome. The *eps* loci are numerous (*eps1, eps2, dsrO, dsrV, levO, gtf, it3, it4* and often divergent from one genome to another. This high diversity fully justifies the method chosen to establish an inventory of *eps* genes (genome sequencing). Genes of interest were identified on the basis of sequence homology, as proposed in other studies [39]. Though the matrix genes blasted in our study are much more numerous (82 reference genes instead of one single gene of priming glycosyltransferase), the existence of genetic determinants with widely differing sequence cannot completely be excluded. However, we found a large number of genes potentially involved in the production of EPS, whose presence is generally relatively well correlated with the observed phenotypes. This suggests that the majority of genes of interest were identified. It appeared that the strains that induced medium ropiness all display *gff* and produce β-glucan. They represent 10% of the strains in the collection studied, while previous work reported a 22% prevalence for *gff* [14]. The strains that produce β-fructan in the presence of sucrose all exhibit a non truncated levansucrase gene, *levO*. The prevalence of *levO* is 26%, with levan production in 77% of the *levO*-strains. Regarding dextran synthesis and dextranucrase gene (*dsrO*), the relationship between genotype and phenotype is less clear. Indeed, the presence of functional genes is not always sufficient to explain the observed phenotypes. Gene expression and activity of *Dsro* could be modulated by certain environmental factors or the physiological state of cells. In previous studies, we observed that glucan and fructan production from sucrose was not detectable in MRS medium but only in semi defined one [15–16]. Anyway, the glycoside-hydrolases of *O. oeni* are not original as regards both the protein primary structure and the structure of the polymers produced. All the encapsulated *O. oeni* strains displayed a cluster *eps2* which encodes the proteins necessary for reconstituting a wzy-dependent pathway. The absence or the significant truncation of cluster *eps2* are always associated with the absence of the polysaccharidic capsule. Nevertheless, the fact that the strain BAA-1163 is encapsulated, although its *eps2* cluster lacks the priming glycosyltransferase, suggests that internal complementation for priming glycosyltransferase is possible (for example by means of genes *woal* or *it3*). In all cases examined, the capsular polymer contains glucose, galactose and rhamnose. This close monomer composition contrasts with the vast diversity of *eps2* cluster sequences. Differences in the osidic bounds encountered in the repeating unit could still exist, and further structure analyses will be necessary to establish a link between the transferases and the monomers present.

The role of cluster *eps1* and of the isolated genes *it3* and *it4* could not be determined in this study. The advantage of the presence of two *eps* clusters remains obscure, but it is clear that this is a common feature to all genomes in the species. Moreover, this is also the case for *O. kilharae*, the other species in the genus *Oenococcus* [40]. Analysis of conserved domains did not enable to clearly predict the function of the Wzy protein encoded in *eps1* (polymerase or ligase). If Wzy is a polymerase, then *eps1* operon would direct the synthesis of an exopolysaccharide. The wzy-dependent synthesis route would be duplicated (one being encoded by *eps1* and the other by *eps2*) with production of two distinct polysaccharide structures, as described for other lactic acid bacteria [41–42]. On the other hand, if the wzy gene in *eps1* encodes a ligase (*WaaL*), the cluster *eps1* may direct the synthesis of an oligosaccharide wherein the ligase then fixes a polysaccharide synthesized by proteins encoded in another cluster (*eps2* for example), on the model of lipopolysaccharide of Gram-negative bacteria [43–44]. In both cases, the product whose synthesis is directed by the *eps1* should be minor because (i) glucuronic acid and phosphoglycerol are never found in the structural analysis of the EPS examined (either soluble or capsular), and (ii) the strains lacking *eps2* cluster but displaying *eps1* show no capsule and produce very low level of soluble EPS in SMD-Glucose.

The distribution of the *eps* genes on the phylogenetic tree is complex. Some genes have clearly been acquired by horizontal transfer after the attack of a bacteriophage (*it4, gtf, dsrV*), while others, could have been acquired earlier in the history of the species (*levO, dsrO, eps1*) or could result of very numerous chromosome modifications (*eps2*). The *eps2* clusters are the most polymorphic among the studied loci. Such a diversity (15 cluster models for 50 genomes) is surprising in a non-pathogenic bacterium as it resembles what is described in *Streptococcus pneumoniae*, in which, *eps* clusters direct the synthesis of a major virulence factor, the pneumococcal capsule [45]. Regarding the cluster organization, the *eps2* clusters, inserted between *amuO* and *recP* also strongly resemble those described for streptococci, whether *S. thermophilus*, in which the *eps* loci are inserted between genes *dxbD* and *pgm*, or *S. pneumoniae*, in which *eps* loci are inserted between genes *dxbB* and *alid* [46–47] or for *Lactococcus* or *Lactobacilli* [48–50]. Genes *dxb* and *alid* are spaced by 10 to 30 kb maximum [47], while *amuO* and *recP* and genes can be distant from 50 kb. This region is the most heterogeneous in the *O. oeni* chromosome [51]. According to Gohluchik et al. [32], the acquisition of *eps* cluster may be accompanied by a large number of changes, spread all along the chromosome. The acquisition of the *eps2* could thus be the cause of the divergence of certain genomes. Loss of cluster *eps2* is rare and in some cases, it is accompanied by the acquisition of the *gff* gene (Champagne strains). The presence of a truncated *eps2* could have been a selection pressure for the stabilization of *gff* (phage remnant). This situation reminds again, what is described in *S. pneumoniae* Type 37 [53].

The fact that the 50 genomes studied possess genes dedicated to EPS metabolism suggests that these polymers are very important for the adaptation of *O. oeni* to its ecological niche. This is even more true for *eps* clusters, not only because they occupy a significant portion of the *O. oeni* small chromosome, but also because the biosynthetic pathway encoded (wzy dependent) is energy consuming [9,54–56]. It is generally claimed that capsular polysaccharides have a mainly protective role while free EPS are interesting from a technological point of view [49,57]. The production of soluble polysaccharides by the strains studied is low in the absence of sucrose (<80 mg/L), but similar to that
described for some other lactic acid bacteria [14,16,49,55–56], or for \textit{O. oeni} in wine [13]. Thirty-two out of 43 strains examined are encapsulated (75\%), against 30\% for \textit{S. thermophilus} [57] or 50\% for \textit{S. pneumoniae} [47]. In \textit{S. pneumoniae}, the capsule is an essential virulence factor. The capsule could thus be a key element for \textit{O. oeni} survival in the host environment. In general, capsular EPS do not constitute an energy supply for the cell that produces them [58–59]. These should rather constitute a protective layer against desiccation, osmotic acid or cold stress, digestion by lysozyme, or against toxic compounds such as alcohol or sulphur dioxide [50,60–63]. EPS could also play a role in biofilm formation, thereby facilitating the colonization of various ecosystems and especially grapes pellicules, barrels and other wine-making material [14,44,59,64–66]. As regards the protection against phage attacks, opposite effects have been described: certain EPS are specifically recognized by certain phages and predispose bacteria to the attack by these phages, while others would be a protective barrier [57,67]. It might be interesting in the future to connect the diversification of \textit{eps} genes with the high variability in Oenophages recently described [12,68,69].

### Supporting Information

Table S1 \textit{In silico} inventory of \textit{eps} genes. List of \textit{eps} genes encountered in the initial database and then, in the 50 genome sequences studied, \textit{locus} by \textit{locus} \textit{eps}1 and \textit{eps}2 clusters, isolated glycosyltransferase and glycoside hydrolase genes, and genes involved in precursor synthesis.

(XLSX)

### Author Contributions

Conceived and designed the experiments: TD CM MRS MDL. Performed the experiments: MD MV MF CMS PL MR PW MP MDL. Analyzed the data: MD HCS PL MDL. Contributed reagents/materials/analysis tools: JC VM EG. Wrote the paper: MD MDL.

### References


