Genetic diversity and population structure of *Saccharomyces cerevisiae* isolated from Turkish sourdough by iPBS-retrotransposons markers

Furkan Aydin1 · Göksel Özer2 · Mehtap Alkan2 · İbrahim Çakır3

Received: 5 May 2022 / Revised: 30 September 2022 / Accepted: 29 October 2022 / Published online: 8 November 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Molecular DNA markers are valuable tools for analyzing genetic variation among yeast from different populations to reveal the genetically different autochthonous strains. In this study, we employed inter-primer binding site (iPBS) retrotransposon polymorphism to assess the genetic variation and population structure of 96 *Saccharomyces cerevisiae* isolates from four different regions in Turkey. The nine selected iPBS primers amplified 102 reproducible and scorable bands, of which 95.10% were polymorphic with an average of 10.78 polymorphic fragments per primer. The average polymorphism information content and the resolving power were 0.26–3.58, respectively. Analysis of molecular variance (AMOVA) revealed significant (*P* < 0.001) genetic differences within populations (88%) and between populations (12%). The unweighted pair group mean with arithmetic (UPGMA) dendrogram grouped 96 *S. cerevisiae* strains into two main clusters, where the highest probability of the data elucidating the population structure was obtained at Δ*K* = 2. There was not an obvious genetic discrimination of the populations according to geographical regions on UPGMA, supported by principal coordinate analysis. However, the individuals of the closer provinces in each population were more likely to group together or closely. The results indicate that iPBS polymorphism is a useful tool to reveal the genetically diverse autochthonous *S. cerevisiae* strains that may be important for the production of sourdough or baked goods.

Keywords Sourdough · Yeast · iPBS · Retrotransposons · Marker · Heterogeneity

Introduction

Sourdough is one of the oldest starters to produce baked goods and cereal-based fermented food products. It enhances the end product’s sensory, rheology, and shelf-life due to its complex microbiota composed of endogenous lactic acid bacteria (LAB) and yeast (Pino et al. 2022). The most significant roles of this microbiota are acidification, flavor formation, and leavening of the dough, where yeast produces aroma compounds and CO₂ (De Vuyst et al. 2016). Many yeast species have been identified from sourdough worldwide. In particular, *Saccharomyces cerevisiae* is the dominant species in sourdough ecosystems (Boyaci-Gunduz and Erten, 2020; Korcari et al. 2020; Pino et al. 2022; Valmorri et al. 2010; Yang et al. 2020). It produces CO₂ as a result of fermentation, resulting in dough expending. It also contributes to developing and strengthening the gluten network in the dough, thus enhancing the texture of the bread (Verheyen et al. 2014) and improving the flavor by secreting aroma compounds, such as aldehydes and esters (Birch et al. 2013).

The technological differences among yeast strains depend on their intraspecific genetic diversity (Huys et al. 2013), which are strongly stimulated by the geographical and ecological differences (Liu et al. 2021; Martínez et al. 2007; Wang et al. 2012). The heterogeneity between the *S. cerevisiae* populations may provide different functional traits for sourdough yeast (Palla et al. 2020) and was previously attributed to different technological characteristics (Drumonde-Neves et al. 2018). Molecular markers are valuable tools to assess this heterogeneity and analyze the population structure of yeast from different geographical...
origins. Up to now, several molecular markers such as inter-simple sequence Repeat (ISSR), randomly amplified polymorphic DNA (RAPD), simple sequence repeat (SSR), and single-nucleotide polymorphisms (SNPs) have been employed to evaluate the genetic variation and population structure of \textit{S. cerevisiae} strains from sourdough and other food matrices (Gallego et al. 2005; Legras et al. 2014; Liu et al. 2021; Schacherer et al. 2009; Schuller et al. 2012; Palla et al. 2020; Yang et al. 2020). Each marker system has advantages and disadvantages, and the selection depends on the species, available equipment, technical expertise, and research funding (Yeken et al. 2022).

Retrotransposons are mobile genetic components distributed throughout the eucaryote genome and divided into two main groups: the long terminal repeat (LTR) retrotransposons and the non-LTR retrotransposons. They copy themselves and jump to another region while the original copy stays on its original locus, increasing the genome size and leading to variations (Kalander et al. 2011). Their integration constructs new joints between the genomic DNA and their conserved ends, making retrotransposons a perfect tool as targeting sites for molecular markers (Kalander and Schulman 2006; Muszewska et al. 2011). Studies revealed the success of inter-LTR PCR fingerprinting in assessing the genetic diversity in yeast previously (Fasoli et al. 2016; Sohier et al. 2009). Retrotransposons have a diverse pattern distribution within the individuals of \textit{S. cerevisiae}, making strain differentiation (Beau REGARD et al. 2008; Bleykasten-Grosshans et al. 2013). Most of the retrotransposon-based methods, such as retrotransposon-microsatellite amplified polymorphism (REMAP), inter-retrotransposon amplified polymorphism (IRAP), and inter-LTR polymorphism requires the sequence information for primer designing, making their use species-specific. Kalendar et al. (2010) introduced a universal retrotransposon-based DNA marker technique, inter-Primer Binding Site (iPBS) retrotransposons, without requiring prior sequence information. Aydin et al. (2020) previously reported the utility of iPBS polymorphism in the yeast genome. The genetic variation in \textit{S. cerevisiae} individuals was notably high, making this PCR-based marker method an excellent candidate to evaluate genetic differentiation and the population structure of \textit{S. cerevisiae}. iPBS retrotransposon markers have not been previously used to investigate the genetic variation and the population structure of \textit{S. cerevisiae} from sourdough worldwide. This study aims to assess the genetic variation and population structure of \textit{S. cerevisiae} strains from different geographic origins (Central Anatolia region, Black Sea region, Mediterranean region, and Aegean region) by iPBS retrotransposon marker system.

### Materials and methods

#### Yeast isolates

Ninety-six \textit{S. cerevisiae} strains isolated from Type I sourdough samples were supplied from Bolu Abant Izzet Baysal University Food Microbiology Laboratory Culture Bank. The samples belonged to Central Anatolia region (CAR; \( n = 8 \)), Black Sea region (BSR; \( n = 8 \)), Mediterranean region (MER; \( n = 8 \)), and Aegean region (AER; \( n = 8 \)). Three \textit{S. cerevisiae} strains were selected for population analyses for each sample. The detailed information regarding the isolates and the isolation region is given in Table 1. In addition, the sampling locations of sourdough from which the isolates were obtained is given in Fig. 1.

#### DNA extraction

DNA was extracted using DNeasy Blood and Tissue kit (Qiagen, Cat No./ID: 69504, Hilden, Germany) according to manufacturer’s instructions. The final quantity of the resultant DNA was determined using the DS-11 FX+ spectrophotometer (Denovix, USA) and diluted to 50 ng/μL with sterile ultra-pure water.

#### Molecular confirmation

The molecular confirmation of \textit{S. cerevisiae} isolates was made using species-specific primer pairs SC1 (5’ – AAC GGT GAG AGA TTT CTG TGC – 3’) and SC2 (5’ – AGC TGG CAG TAT FCC CAC AG – 3’) designed by Josepa et al. (2000). As a positive control, \textit{S. cerevisiae} strain 105-E2 deposited in GenBank under MK358173 accession number was used. The PCR reactions were performed with a 50 μL reaction mixture containing 1x PCR reaction buffer, 0.4 μM of SC1 and SC2 primers, 200 μM of each dNTPs, 20 ng template DNA, and 1.5-unit Ampliqon TEMPase Hot Start DNA polymerase (Berntsen, Rdovre, Denmark). The PCR amplification consisted of 5 min initial denaturation at 94 °C, followed by 30 cycles with 30 s denaturation at 94 °C, 30 s annealing at 50 °C and 1 min extension at 72 °C and 7 min final extension at 72 °C. The amplicons were analyzed on 1.4% agarose gel, stained with ethidium bromide, and visualized using a gel imaging system (G:BOX F3, Syngene, Cambridge, UK) under UV.

#### iPBS analyses

Three randomly selected strains were initially screened using 83 iPBS primers designed by Kalendar et al. (2010). As a result of preliminary studies, nine primers produced clearer
and sharper fragments (Table 2). The reaction mix included 50 ng template DNA, 1x Dream Taq Buffer, 1 μM of primer, 200 μM of dNTPs, 1.2-unit of Dream Taq DNA polymerase (Thermo Fischer Scientific, Waltham, MA, USA) and 0.04-unit of Pfu DNA polymerase (Thermo Fischer Scientific, Waltham, MA, USA). The amplification was performed using a T100 thermocycler. The reaction included initial denaturation at 94 °C for 3 min, followed by 35 cycles with 30 s denaturation at 94 °C, 30 s annealing at 50–65 °C depending on primer, 2 min extension at 72 °C and ending with one cycle of final extension at 72 °C for 10 min. The amplified fragments were analyzed on 1.6% (w/v) agarose gel with TAE, stained with ethidium bromide and visualized using a gel imaging system (G:BOX F3, Syngene, Cambridge, UK) under UV.

**Evaluation of iPBS data**

Strong, clear, and unambiguous PCR bands were scored manually as present (1) and absent (0) at their positions to build a binary data matrix. We used a 100 bp DNA ladder (Solis BioDyne, Tartu, Estonia) from each end of the gel to validate the gel-to-gel normalization of migrations. The discriminatory power of each iPBS primer was assessed by determining

| Sampling region | Province | Strain code |
|-----------------|----------|-------------|
| **Central Anatolia (n = 8)** | Aksaray (n = 1) | 1, 2, 3 |
| | Ankara (n = 3) | 198, 199, 200, 203, 204, 205, 208, 209, 210 |
| | Karaman (n = 2) | 23, 24, 25, 28, 29, 30 |
| | Konya (n = 1) | 103, 104, 105 |
| | Sivas (n = 1) | 228, 229, 230 |
| **Black Sea (n = 8)** | Amasya (n = 1) | 254, 255, 256 |
| | Bolu (n = 1) | 5, 6, 7 |
| | Giresun (n = 1) | 153, 154, 155 |
| | Samsun (n = 2) | 223, 224, 225, 248, 249, 250 |
| | Trabzon (n = 3) | 238, 239, 240, 243, 244, 245, 288, 289, 290 |
| **Mediterranean (n = 8)** | Antalya (n = 3) | 108, 109, 110, 113, 114, 115, 118, 119, 120 |
| | Adana (n = 3) | 188, 189, 190, 278, 279, 280, 283, 284, 285 |
| | Mersin (n = 2) | 14, 15, S8, 16, 17, 18 |
| **Aegean (n=8)** | Aydin (n = 1) | 293, 294, 295 |
| | Balikesir (n = 1) | 193, 194, 195 |
| | Denizli (n = 1) | 273, 274, 275 |
| | Izmir (n = 2) | 39, 40, 41, 44, 45, 46 |
| | Manisa (n = 2) | 168, 169, 170, 218, 219, 220 |
| | Muğla (n = 1) | 35, 36, 37 |

*n = Number of samplings*

---

Fig. 1 Location of sourdough samples from which the *S. cerevisiae* strains were isolated. Different colors represent different sampling regions. CAR: Central Anatolia region, BSR black sea region, MER Mediterranean region, AER aegean region.
the resolving power (RP) and the polymorphic information content (PIC), as proposed by Prevost and Wilkinson (1999) and Roldán-Ruiz et al. (2000), respectively. The PIC value of each iPBS primer was determined using the formula \( \text{PIC} = 2f(1 − f) \); \( f \) is the frequency of the amplified allele. For the determination of the RP value for a dominant single primer amplification reaction marker, the band informativeness \( (I_b) \) was calculated first, as follows: \( I_b = 1 – (2 \times 0.5 - p) \), where \( p \) represents the proportion of yeast isolates. Then RP was calculated as the sum of \( I_b \) for each band obtained for each primer.

The binary data was converted into a genetic similarity matrix using Jaccard’s similarity coefficient based on all primers’ proportion of shared alleles and Unweighted Pair Group Mean with Arithmetic (UPGMA) analysis was estimated with the vegan library Ver. Vegan 2.4.4 in R (Oksanen et al. 2017; R Studio Team 2020). Shannon’s information index, Nei’s gene diversity, number of effective and different alleles, analysis of molecular variance (AMOVA), and principal principal coordinate analysis (PCoA) were analyzed using the software program GenAlEx 6.5 (Peakall and Smouse 2012).

The software STRUCTURE (v.2.3.4) was used to determine the population structure of the 96 \( S. \) cerevisiae strains. The number of hypothetical subpopulations \( (K) \) was estimated with 10 independent values \( (K \) from 1 to 10). The length of burning and Markov Chain Monte Carlo (MCMC) repeats after burning were specified as 100,000 for each. STRUCTURE HARVESTER was used to test the best \( \Delta K \) value using Evanno method (Earl 2012; Evanno et al. 2005; Pritchard et al. 2020).

**Results and discussion**

The nine selected iPBS primers amplified 102 reproducible and scorable bands, of which 95.10% were polymorphic. The iPBS primers used in this study had higher polymorphism rates than those obtained by Liu et al. (2021), who used ISSR primers for analyzing the genetic variation within industrial brewing yeast. The number of amplified fragments per iPBS primer ranged from 7 (2270) to 16 (2395), giving a ratio of 10.78 fragments per primer (Table 2). PIC and RP values were determined to reveal the polymorphism rate and ability to differentiate the genotypes. The average RP value, indicating the discriminatory power of the selected primers, was 3.58, which ranged from 2.06 (2270) to 4.77 (2078). On the other hand, the iPBS 2078 yielded the highest PIC value (0.31), while 2376 iPBS primer gave the lowest value (0.20). The average PIC value calculated was 0.26. These values confirm the utility of the selected iPBS primers for assessing the genetic diversity and grouping genetically diverse \( S. \) cerevisiae strains. Aydin et al. (2020) calculated the PIC value as 0.20-0.30 for different yeast species using eight iPBS primers, which is in accordance with our results. Compared to other dominant markers, Lathar et al. (2010) and Liu et al. (2021) reported slightly higher PIC values for \( S. \) cerevisiae strains using RAPD and ISSR primers, respectively.

The mean values for some genetic diversity indices, which are the number of alleles (1.29 ± 0.03), the effective number of different alleles, analysis of molecular variance (AMOVA), and principle principal coordinate analysis (PCoA) were analyzed using the software program GenAlEx 6.5 (Peakall and Smouse 2012).

**Table 2** The information of iPBS markers to evaluate the genetic diversity of yeast

| Primer IDs | Primer sequence (5’–3’) | Ta (°C) | GC (%) | TB | PB | PIC | RP |
|------------|-------------------------|--------|--------|----|----|-----|----|
| 2080       | CAGACGGGCAGCCA          | 63     | 75.0   | 13 | 13 | 0.29 | 4.68|
| 2395       | TCCCCCGGCGGATCGCCA      | 53     | 72.2   | 16 | 15 | 0.24 | 4.66|
| 2078       | GCCGAGTCCGCA            | 62     | 75.0   | 12 | 11 | 0.31 | 4.77|
| 2242       | GCCCCCATGTGCGGCCA       | 57     | 77.8   | 11 | 11 | 0.29 | 3.99|
| 2376       | TAGATGGCAGCCA           | 52     | 50.0   | 10 | 9  | 0.20 | 2.40|
| 2271       | GGTCTGGGTGCCA           | 60     | 69.2   | 13 | 12 | 0.21 | 3.38|
| 2386       | CTGATCGAACC            | 50     | 50.0   | 11 | 10 | 0.30 | 3.55|
| 2076       | GCTCGGTGCCA            | 59     | 66.7   | 9  | 9  | 0.24 | 2.77|
| 2270       | ACCTGGCGTGCCA          | 65     | 69.2   | 7  | 7  | 0.27 | 2.06|

| Total       | 102                     | 97     | 0.26   | 3.58|
| Avg./primer | 11.33                   | 10.78  |        |     |

\( Ta \) annealing temperature, \( GC \) guanine-cytosine content, \( TB \) total band, \( PB \) polymorphic band, \( PIC \) polymorphism information content, \( RP \) resolving power
Table 3 Some parameters to evaluate the genetic diversity among populations by iPBS primers

| Population* | Na** | Ne | I | h |
|-------------|------|----|---|---|
| CAR (n=24)  | 1.37 ± 0.06 | 1.26 ± 0.03 | 0.24 ± 0.03 | 0.16 ± 0.02 |
| BSR (n=24)  | 1.24 ± 0.07 | 1.22 ± 0.03 | 0.20 ± 0.03 | 0.13 ± 0.02 |
| MER (n=24)  | 1.26 ± 0.07 | 1.20 ± 0.03 | 0.19 ± 0.02 | 0.12 ± 0.02 |
| AER (n=24)  | 1.30 ± 0.07 | 1.23 ± 0.03 | 0.21 ± 0.03 | 0.14 ± 0.02 |
| Overall (n=96) | 1.29 ± 0.03 | 1.23 ± 0.02 | 0.21 ± 0.01 | 0.14 ± 0.01 |

*CAR central Anatolia region, BSR black sea region, MER Mediterranean region, AER aegean region
**Na: The number of alleles. Ne: The effective number of alleles. I: Shannon’s information index. h: Nei’s (1973) gene diversity

Analysis of molecular variance (AMOVA) was conducted considering within and among the populations. Based on the AMOVA results, significant (P < 0.001) genetic differences were obtained within (88%) and among populations (12%) (Table 4). It indicated that the genetic variation was mostly among the individuals, rather than populations. The higher genetic difference within the populations addresses greater levels of subdivision and hierarchy (Barut et al. 2020). These results are firmly in accordance with those reported by the other researchers using microsatellite markers (Bigey et al. 2021; Borlin et al. 2016; Muller and McCusker 2009; Schuller et al. 2012). It should be noted that a significant genetic diversity is present within the individuals of the populations, which can provide beneficial information for screening technologically superior autochthonous strains for the production of sourdough or baked goods. Retrotransposons compromise 3.35% of the *Saccharomyces cerevisiae* genome (Carr et al. 2012) and have been reported to possess different pattern distribution within individuals, contributing to the strains’ genetic and technological diversity (Beauregard et al. 2008; Bleykasten-Grosshans et al. 2013).

The UPGMA dendrogram grouped 96 *S. cerevisiae* strains into two main clusters, where three sub-clusters were formed under each cluster (Fig. 2). Cluster I compromised 58 isolates, while Cluster II included the rest. The Bayesian clustering model conducted by STRUCTURE divided 96 *S. cerevisiae* isolates into two populations, as well. The highest probability of the data elucidating the population structure was obtained at ΔK=2 (Fig. 3) which utilizes the genetic data to evaluate the population membership without predefined populations and assigns the individuals to clusters based on multilocus genotypes (Chen et al. 2007). The genetic variation in Cluster I (F_{ST}=0.53) was higher than that of Cluster II (F_{ST}=0.33) due to grouping genetically more diverse individuals in Cluster I. The population structure of the individuals is also given in Fig. 4, which is in accordance with the clusters obtained by UPGMA.

There was not an obvious genetic discrimination of the populations according to geographical regions on UPGMA and PCoA. However, the same populations’ individuals had a tendency to group together or closely, such as Denizli and Izmir or Amasya and Trabzon provinces. The backslopping of the sourdough occurs very frequently depending on the bread production frequency. The polymorphism in yeast genome occurs during the vegetative phase where the meiosis is a rare. If the yeast reproduces constantly adapting its own environment, the genetic variation occurs according to ecological and geographic conditions (Martínez et al. 2007). In addition, the individuals of the same populations were also distributed in different clusters, confirmed by the PCoA (Fig. 5). It defined two major groups and depicted the genetic relationship between the individuals. Similar results where the close grouping of the geographically related individuals was also reported for *S. cerevisiae* using different DNA markers by several authors (Capece et al. 2016; Jeyaram kiy et al. 2014). Up to date, the population structure of *S.

Table 4 Results of AMOVA of *Saccharomyces cerevisiae* by the cluster algorithm

| Source          | d.f* | SS     | MS     | Est. Var | %  | F_{ST} | P       |
|-----------------|------|--------|--------|----------|----|--------|---------|
| Among populations | 3    | 106.135| 35.378 | 1.133    | 12 | 0.122  | < 0.001 |
| Within populations | 92   | 752.208| 8.176  | 8.176    | 88 |        |         |
| Total           | 95   | 858.344| 9.310  |          |    |        |         |

*d.f* degrees of freedom, SS sums of squares, MS mean square, Est. var. estimated variation, F_{ST} f statistics. P = significance level
The findings of this study suggest that the iPBS markers also provide helpful information regarding the genetic variation and the population structure of *S. cerevisiae* from different locations. Similarly, the utility of iPBS retrotransposons in fungi for the assessment of genetic diversity and population structure was also reported by other researchers (Ates et al. 2019; Erper et al. 2021; Ozer et al. 2016, 2017).

In conclusion, the genetic variation and the population structure of 96 *S. cerevisiae* strains were investigated using iPBS markers. To the best of authors’ knowledge, it is the first study to determine the genetic diversity and population structure of *S. cerevisiae* using iPBS markers worldwide. The findings revealed that iPBS markers were useful in assessing the genetic diversity of autochthonous *S. cerevisiae* strains from different locations. This technique has a strong potential to reveal the genetically diverse *S. cerevisiae* strains that may later be correlated with some technological characteristics necessary for producing sour-dough and baked goods. Therefore, it may be a useful tool to provide functional starters by revealing the heterogeneity. To this aim, more studies are required to investigate the relationship between technological characteristics and polymorphism obtained by iPBS markers.
Fig. 4  Population structure of the individuals at $\Delta K = 2$ based on the iPBS-retrotransposons data. CAR central Anatolia region, BSR black Sea region, MER mediterranean region, AER aegean region

Fig. 5  The distribution of *Saccharomyces cerevisiae* populations according to Principle Coordinate Analysis. CAR central Anatolia region, BSR black Sea region, MER mediterranean region, AER aegean region
### Table 5  Nei’s genetic similarity matrix of *Saccharomyces cerevisiae* populations

| Populations* | CAR | BSR | MER | AER |
|--------------|-----|-----|-----|-----|
| CAR          | 1   | 0.972 | 0.985 | 0.978 |
| BSR          | 1   | 0.972 | 0.965 | 0.967 |
| MER          | 0.985 | 0.965 | 1 | 0.963 |
| AER          | 0.978 | 0.967 | 0.963 | 1 |

*CAR central anatolia region, BSR black Sea region, MER mediterranean region, AER aegan region*

### Author contributions
Design of the study: FA, GÖ, İÇ. Supply of the strains: FA, İÇ. Methodology: FA, MA, GÖ. Statistical analyses: FA, MA, GÖ. Writing and editing: FA, GÖ, İÇ. Supervision: GÖ, İÇ. All authors have read and agreed to the published version of the manuscript.

### Funding
The authors would like to thank the Scientific and Technological Research Council of Turkey (TUBITAK) for its financial support under project code 1210580.

### Data availability statement
All the data generated or analyzed during this study are included in this published article.

### Declarations

#### Conflict of interest
The authors declare that there is no conflict of interest for this research.

#### Authorship policies
All authors conform to the Archive of Microbiology’s authorship policies.

### References

Ates D, Altinok HH, Ozkuru E, Ferik F, Erdoganus S, Can C, Tanyolac MB (2019) Population structure and linkage disequilibrium in a large collection of *Fusarium oxysporum* strains analysed through iPBS markers. J Phytopathol 167(10):576–590. https://doi.org/10.1111/jph.12848

Aydin F, Ozert, Alkan M, Cakir I (2020) The utility of iPBS retrotransposons markers to analyze genetic variation in yeast. Int J Food Microbiol 325:108647. https://doi.org/10.1016/j.ijfoodmicro.2020.108647

Barut M, Nadeem MA, Karakoy T, Baloch FS (2020) DNA fingerprinting and genetic diversity analysis of world quinoa germplasm using iPBS-retrotransposon marker system. Turk J Agric For 44(5):479–491. https://doi.org/10.3906/tar-2001-10

Beauregard A, Curcio MJ, Belfort M (2008) The take and give between retrotransposable elements and their hosts. Annu Rev Genet 42:587–617. https://doi.org/10.1146/annurev.genet.42.110807.091549

Bigey F, Segond D, Friedrich A, Guze nec S, Bourgais A, Huyge L, Agier N, Nidelet T, Sicard D (2021) Evidence for two main domestication trajectories in *Saccharomyces cerevisiae* linked to distinct bread-baking processes. Curr Biol 31(4):722–732. https://doi.org/10.1016/j.cub.2020.11.016

Birch AN, Petersen MA, Arneborg N, Hansen AS (2013) Influence of commercial baker’s yeasts on bread aroma profiles. Food Res Int 52(1):160–166. https://doi.org/10.1016/j.foodres.2013.03.011

Bleykasten-Grosshans C, Friedrich A, Schacherer J (2013) Genomewide analysis of intraspecific transposon diversity in yeast. BMC Genom 14(1):1–13. https://doi.org/10.1186/1471-2164-13-399

Bozorgzadeh I (2016) Cell-associ-ated *Saccharomyces cerevisiae* population structure revealed high-level diversity and perennial persistence at sartournes wine estates. Appl Envir Microbiol 82(10):2909–2918. https://doi.org/10.1128/AEM.03627-15

Boyaci-Gunduz CP, Erten H (2020) Predominant yeasts in the sour-doughs collected from some parts of Turkey. Yeast 37(9–10):449–466. https://doi.org/10.1002/yea.3500

Capece A, Granchi L, Guerrini S, Mangani S, Romaniello R, Vincen- zini M, Romano P (2016) Diversity of *Saccharomyces cerevisiae* strains isolated from two Italian wine-producing regions. Front Microbiol 7:1018. https://doi.org/10.3389/fmicb.2016.01018

Carr M, Bensasson D, Bergman CM (2012) Evolutionary genomics of transposable elements in *Saccharomyces cerevisiae*. PLoS ONE 7(11):e50978. https://doi.org/10.1371/journal.pone.0050978

Chen C, Durand E, Forbes F, Francois O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. Mol Ecol Notes 7(5):747–756. https://doi.org/10.1111/j.1471-2286.2007.01769.x

De Vuyst L, Hartth H, Van Kerrebroeck S, Leroy F (2016) Yeast diversity of sourdoughs and associated metabolic properties and functionalities. Int J Food Microbiol 239:26–34. https://doi.org/10.1016/j.ijfoodmicro.2016.07.018

Drumonde-Neves J, Franco-Duarte R, Vieira E, Mendes I, Lima T, Schuller D, Pais C (2018) Differentiation of *Saccharomyces cerevisiae* populations from vineyards of the azores archipelago: geography vs ecology. Food Microbiol 74:151–162. https://doi.org/10.1016/j.fm.2018.03.017

Earl DA (2012) Structure harvester: a website and program for visualizing structure output and implementing the evanno method. Conserv Genet Resour 4(2):359–361. https://doi.org/10.1007/s12686-011-9548-7

Erper I, Ozert G, Kalendar R, Avci S, Yildirim E, Alkan M, Turkkan M (2021) Genetic diversity and pathogenicity of *Rhizoctonia* spp isolates associated with red cabbage in samsun (Turkey). J Fungi 7(4):23. https://doi.org/10.3390/jof7030234

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol 14(8):2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Fasoli G, Barrio E, Tofalo R, Suzzi G, Belloch C (2016) Multilocus analysis reveals large genetic diversity in *Kluyveromyces marxianus* strains isolated from parmigiano reggiano and pecorino de farindola cheeses. Int J Food Microbiol 233:1–10. https://doi.org/10.1016/j.ijfoodmicro.2016.05.028

Gallego FJ, Perez MA, Nuñez Y, Hidalgo P (2005) Comparison ofRAPDs, AFLPs and SSR markers for the genetic analysis of yeast strains of *Saccharomyces cerevisiae*. Food Microbiol 22(6):561–568. https://doi.org/10.1016/j.fm.2004.11.019

Gayevisky V, Klære S, Knight S, Goddard MR (2014) ObStruct: a method to objectively analyse factors driving population structure of individuals using the software structure: a simulation study. Mol Ecol 1(4):2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Gayevskiy V, Klaere S, Friedrich A, Schacherer J (2013) Taxonomy and biodiversity of sourdough yeasts and lactic acid bacteria. In: Gobbetti M, Gänzle M (eds) Handbook on sourdough biotechnology. Springer, US, pp 105–154

Huys G, Daniel H-M, De Vuyst L (2013) Taxonomy and biodiversity of sourdough yeasts and lactic acid bacteria. In: Gobbetti M, Gänzle M (eds) Handbook on sourdough biotechnology. Springer, US, Boston, MA, pp 105–154

Jeyaram K, Tamang JP, Capece A, Romano P (2011) Geographical origins domesticated for rice-based ethnic fermented beverages production in North East India. Antonie Leeuwenhoek 100(4):569–578. https://doi.org/10.1007/s10482-011-9612-z
Josepa S, Guillamon JM, Cano J (2000) PCR differentiation of *Saccharomyces cerevisiae* from *Saccharomyces bayanus*/*Saccharomyces pastorianus* using specific primers. FEMS Microbiol Lett 193(2):255–259. https://doi.org/10.1111/j.1574-6968.2000.tb09433.x

Kalendr R, Schulman AH (2006) IRAP and REMAP for retrotransposon-based genotyping and fingerprinting. Nat Protoc 1(5):2478–2484. https://doi.org/10.1038/nprot.2006.377

Kalendr R, Antonius K, Smykal P, Schulman AH (2010) iPBS: a universal method for DNA fingerprinting and retrotransposon isolation. Theor Appl Genet 121(8):1419–1430. https://doi.org/10.1007/s00122-010-1398-2

Kalendr R, Flavell AJ, Ellis THN, Sjakste T, Moisy C, Schulman AH (2011) Analysis of plant diversity with retrotransposon-based molecular markers. Hereditas 106(4):520–530. https://doi.org/10.1038/hdy.2010.93

Korcari D, Ricci G, Quattrini M, Fortina MG (2020) Microbial consortia involved in fermented spelt sourdoughs: dynamics and characterization of yeasts and lactic acid bacteria. Lett Appl Microbiol 70(1):48–54. https://doi.org/10.1111/lam.13241

Lathar PK, Sharma A, Thakur I (2010) Isolation and random amplified polymorphic DNA (RAPD) analysis of wild yeast species from 17 different fruits. J Yeast Fungal Res 1(8):146–151. https://doi.org/10.5897/JYFR.9000033

Legras JL, Erny C, Charpentier C (2014) Population structure and comparative genome hybridization of European flour yeast reveal a unique group of *Saccharomyces cerevisiae* strains with few gene duplications in their genome. PLoS ONE 9(10):e108089. https://doi.org/10.1371/journal.pone.0108089

Liu J, Li X, Liu Y, Xing C, Xie Y, Cai G, Lu J (2021) Evaluation of genetic diversity and development of core collections of industrial brewing yeast using ISSR markers. Arch Microbiol 203(3):1001–1008. https://doi.org/10.1007/s00203-020-02091-8

Martinez C, Cosgaya P, Vásquez C, Gac S, Ganga A (2007) High degree of correlation between molecular polymorphism and geographic origin of wine yeast strains. J Appl Microbiol 103(6):2185–2195. https://doi.org/10.1111/j.1365-2672.2007.03493.x

Muller LA, McCusker JH (2009) Microsatellite analysis of genetic diversity among clinical and nonclinical *Saccharomyces cerevisiae* isolates suggests heterozygote advantage in clinical environments. Mol 18(13):2779–2786. https://doi.org/10.1111/j.1569-294X.2009.04234.x

Muszewska A, Hoffman-Sommer M, Gryenberg M (2011) LTR retrotransposons in fungi. Plos One 6(12):e29425. https://doi.org/10.1371/journal.pone.0029425

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci U S A 70:3321–3326. https://doi.org/10.1073/pnas.70.12.3321

Oksanen, J, Blanchet, FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O’Hara RB, Simpson GL, Solymos P, Henry M, Stevens H, Szoecs E, Wagner H (2017) Vegan: Community Ecology Package. R package version 2.4–2.2017. <https://cran.r-project.org/web/packages/vegan/index.html> (accessed on 5 April 2022).

Ozer G, Bayraktar H, Baloch FS (2016) iPBS retrotransposons ‘a universal retrotransposons’ now in molecular phylogeny of fungal pathogens. Biochem Syst Ecol 68:142–147. https://doi.org/10.1016/j.bioseb.2016.07.006

Ozer G, Sameeullah M, Bayraktar H, Gore ME (2017) Genetic diversity among phytopathogenic Sclerotiniaceae, based on retrotransposon molecular markers. Phytopathol Mediterr 25–258(10):14601 (/Phytopathol_Mediterr-20379)

Palla M, Cristini C, Giovanetti M, Agnolucci M (2020) Large genetic intraspecific diversity of autochthonous lactic acid bacteria and yeasts isolated from PDO Tuscan bread sourdough. Appl Sci 10(3):1043. https://doi.org/10.3390/app10031043

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update.

Bioinformatics 28:2537–2539. https://doi.org/10.1093/1471-2866.2005.01155.x

Pino A, Russo N, Solieri L, Sola L, Caggia C, Randazzo CL (2012) Microbial consortia involved in traditional sicilian sourdough: characterization of lactic acid bacteria and yeast populations. Microorganisms 10(2):283. https://doi.org/10.3390/microorganisms10020283

Prevost A, Wilkinson MJ (1999) A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. Theor Appl Genet 98(1):107–112. https://doi.org/10.1007/s001220051046

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155(2):945–959. https://doi.org/10.1093/genetics/155.2.945

R Studio Team (2020) RStudio: integrated development for R. PBC, Boston, MA URL http://www.rstudio.com/

Roldán-Ruiz I, Dendaun J, Van Bockstaele E, Depicker A, De Loose M (2000) AFLP markers reveal high polymorphic rates in ryegrasses (Lolium spp.). Mol Breed 6(2):125–134. https://doi.org/10.1023/A:1009680614564

Schacherer J, Shapiro JA, Rudenier DM, Kruglyak L (2009) Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. Nature 458(7236):342–345. https://doi.org/10.1038/nature07670

Schuller D, Cardoso F, Sousa S, Gomes P, Gomes AC, Santos MA, Casal M (2012) Genetic diversity and population structure of *Saccharomyces cerevisiae* strains isolated from different grape varieties and winemaking regions. PLoS ONE 7(2):e32507. https://doi.org/10.1371/journal.pone.0032507

Sohier D, Le Dizes AS, Thuault D, Neveuglise C, Coton E, Casaregola S (2009) Important genetic diversity revealed by inter-LTR PCR fingerprinting of *Kluveromyces marxianus* and *Debaryomyces Hansenii* strains from French traditional cheeses. Dairy Sci Technol 89:569–581. https://doi.org/10.1051/dst/2009032

Tello J, Cordero-Bueso A, Aporta I, Cabellos JM, Arroyo T (2012) Genetic diversity in commercial wineries: effects of the farming system and vinification management on wine yeasts. J Appl Microbiol 112(2):302–315. https://doi.org/10.1111/j.1365-2672.2011.05202.x

Valmorri S, Tofalo R, Settanni L, Corsetti A, Suzzi G (2010) Yeast microbiota associated with spontaneous sourdough fermentations in the production of traditional wheat sourdough breads of the Abruzzo region (Italy). Antonie Leeuwenhoek 97(2):119–129. https://doi.org/10.1007/s10482-009-9392-x

Verheyen C, Jekle M, Becker T (2014) Effects of *Saccharomyces cerevisiae* on the structural kinetics of wheat dough during fermentation. LWT 58(1):194–202. https://doi.org/10.1016/j.lwt.2014.02.050

Wang QM, Liu WQ, Liti G, Wang SA, Bai FY (2012) Surprisingly diverged populations of *Saccharomyces cerevisiae* in natural environments remote from human activity. Mol Ecol 21:5404–5417. https://doi.org/10.1111/j.1365-294X.2012.05732.x

Yang H, Liu T, Zhang G, He G (2020) Intraspecific diversity and fermentative properties of *Saccharomyces cerevisiae* from Chinese traditional sourdough. LWT. https://doi.org/10.1016/j.lwt.2020.109195

Yeken MZ, Emiralioglu O, Ciftci V, Bayraktar H, Palacioglu G, Ozer G (2022) Analysis of genetic diversity among common bean germplasm by start codon targeted (SCoT) markers. Mol Biol Rep. https://doi.org/10.1007/s11033-022-07229-z

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g., a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.