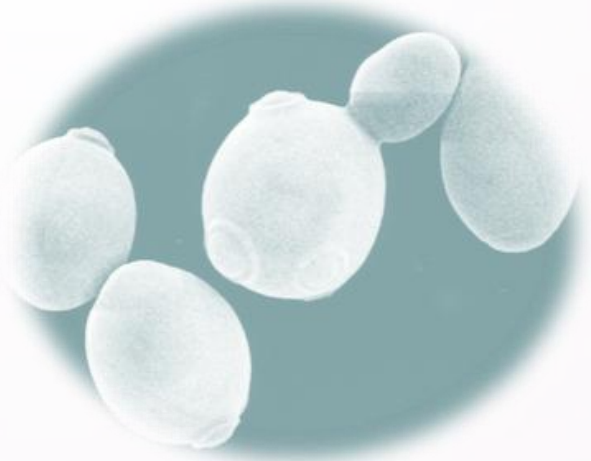
Quality control of *Saccharomyces* yeasts: differentiation of species level and strain grouping using COX 2 gene analysis and MALDI-TOF-MS analysis

Hiroko Kawasaki[1], Shinya Ninomiya[1], Kanae Teramoto[2], Hiroaki Sato[3], Ken-ichiro Suzuki[1]

[1]NITE Biological Resource Center (NBRC), National Institute of Technology and Evaluation (NITE)

[2]JEOL Ltd.,

[3]Resarch Institute for Environmental Management Technology, National Institute of Advanced Industrial Science and Technology (AIST)



Background

Genus *Saccharomyces*

The Genus *Saccharomyces* plays an important role in human activities. They are used as fermenting agents worldwide and stand out as model eukaryotic organisms in various fields of biological science ranging from biochemistry to genomics.

Taxonomy of *Saccharomyces*

The modern phylogenetic concept of *Saccharomyces* based on D1/D2 domain of LSU rDNA, ITS1-5.8S-ITS2, EF-1α, Mt Sm rDNA and COX2, restricts this genus to eight species: *Saccharomyces cerevisiae, S. bayanus*, *S. cariocanus*, *S. kudriavzevii*, *S. mikatae*, *S. paradoxus, S. pastorianus, S. arborilus.* Some of the *Saccharomyces* species were transferred to three genera *Kazachstania*, *Lachancea* and *Naumovia.*

Objective

Saccharomyces **species have various characterizations in strain level**

- **yield of ethanol production**
- **ability for osmotolerance**
- **variety of organic compound etc.**

In NBRC Yeast Collection

- **760 strains of *Saccharomyces* species are collected**
- **551 strains are opened to the public**

Because of the potential value of *Saccharomyces* yeast at the strain level, discrimination and grouping of strains are needed.

In this study, we aimed to apply the high resolution biomarker for grouping and identification of the genus *Saccharomyces.*

High resolution biological marker

LSU rDNA D1/D2 domain:
- main current for identification in species level
- the guideline “representatives of different species tend to show at least 1% divergence in this sequence." by Kurtzman & Robnett (1998)

It is effective for the identification of general ascomycetos yeast, but it is unsuitable for the identification of *Saccharomyces* species.

1. Cytochrome oxidase subunit II (COX2)
- located on mitochondrial gene
- molecular phylogenetic analysis

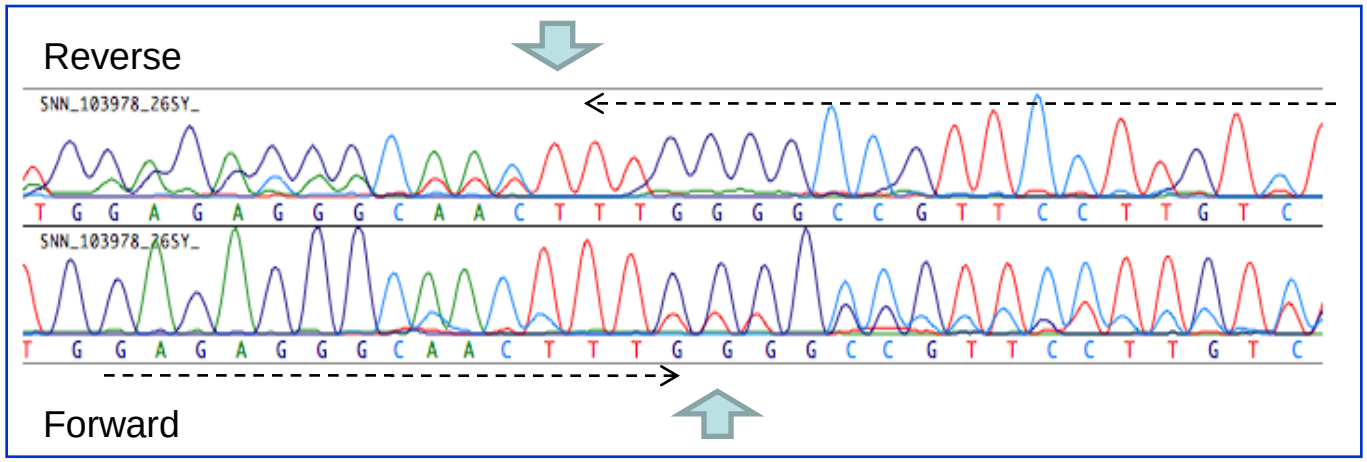
2. Ribosomal subunit proteins
- one component of the cytoplasm ribosome
- comparative mass spectra analysis using MALDI-TOF MS

LSU rDNA D1/D2 Domain vs. COX2 gene

	1	2	3	4	5	6	7	8	9
1. *S. cerevisiae*	–	96	93	98	98	nd	92	93	91
2. *S. cariocanus*	99 (4)	–	94	96	96	nd	91	93	92
3. *S. kudriavzevii*	99 (8)	99 (4)	–	94	94	nd	92	92	88
4. *S. mikatae*	98 (12)	99 (8)	98 (10)	–	98	nd	92	93	90
5. *S. paradoxus*	99 (5)	100 (1)	99 (5)	99 (7)	–	nd	91	92	90
6. *S. arborilus*	98 (10)	99 (3)	99 (5)	99 (9)	99 (4)	–	nd	nd	nd
7. *S. bayanus*	98 (11)	98 (9)	99 (7)	98 (11)	99 (8)	99 (8)	–	97	89
8. *S. pastorianus*	98 (11)	98 (9)	99 (7)	98 (11)	99 (8)	99 (8)	100 (0)	–	90
9. *Lachancea kluyveri*	93 (39)	93	93	92	93	93	93	93	–

The triangle under the left shows similarity (%) of LSU rDNA D1/D2 and the number in parentheses shows the number of base substitution.
The upper right triangle shows similarity (%) of the COX2 gene.
The type strains were used.

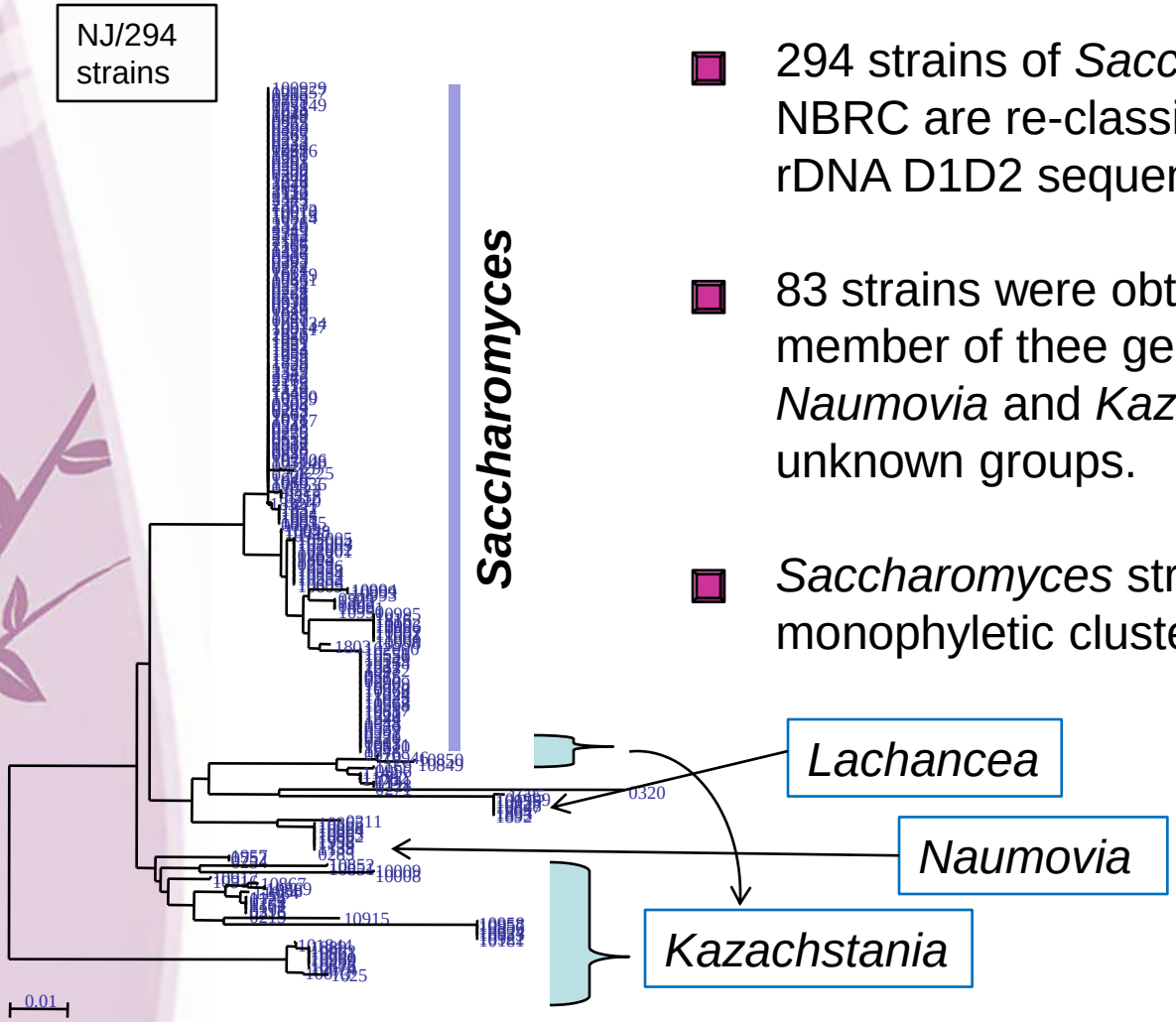
LSU rDNA D1D2 vs COX2 gene



Name	NBRC strain	Sequences	LSU rDNA blast homology search	Cox2 blast homology search	Cluster
Saccharomyces cerevisiae	103978	Triple 'T' insertion at position 106 bases	100% (585/585) Saccharomyces sp. AJ938046 CBS 8615	97% (597/611) Saccharomyces cerevisiae AY244992 CBS 1171T	S. cerevisiae cluster 1

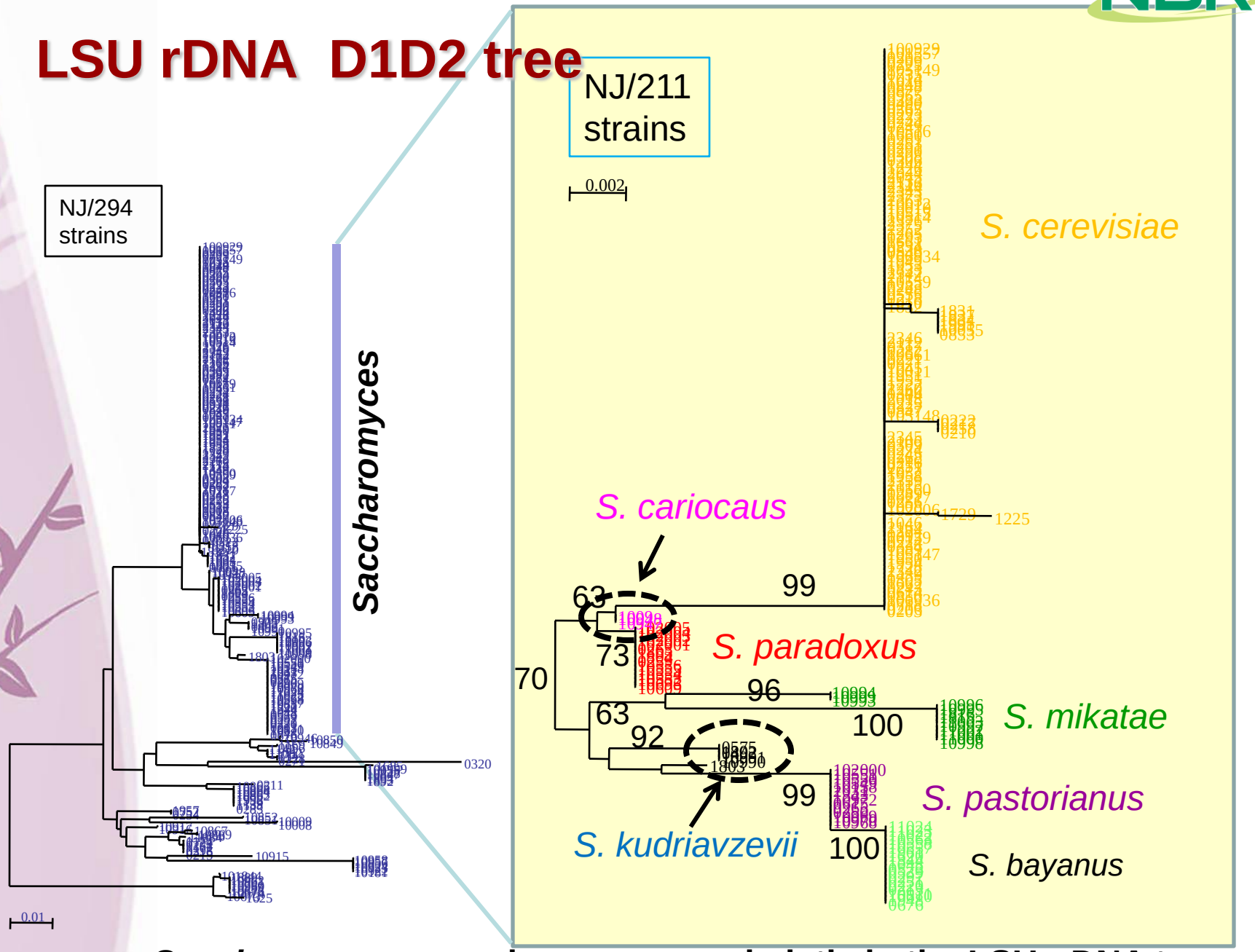
All of the COX2 gene studied are single sequences.

LSU rDNA D1D2 tree



- 294 strains of *Saccharomyces* stocked in NBRC are re-classified based on the LSU rDNA D1D2 sequences.
- 83 strains were obtained to become a member of thee genera *Lachancea, Naumovia* and *Kazachstania and* two unknown groups.
- *Saccharomyces* strains form a monophyletic cluster.

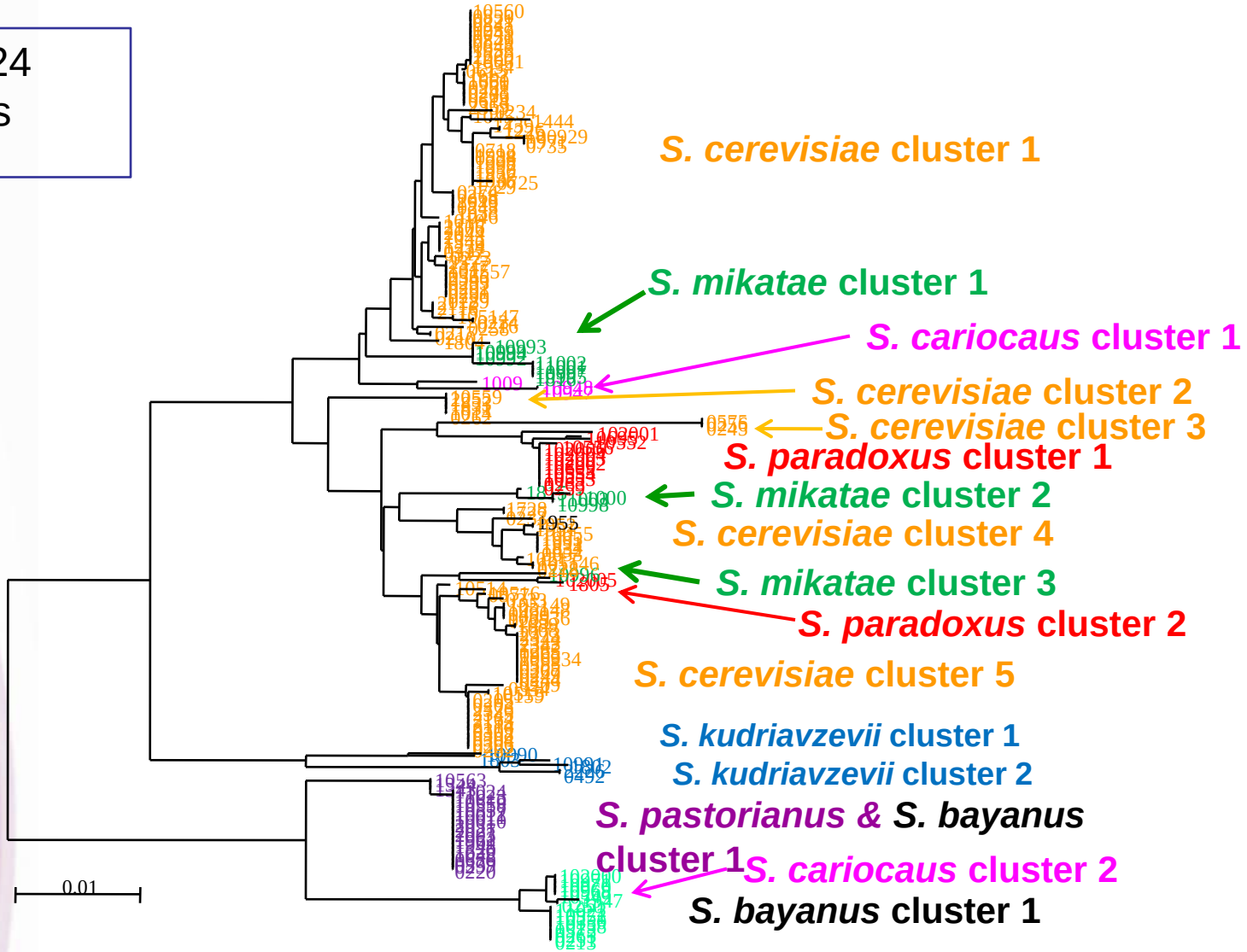
LSU rDNA D1D2 tree



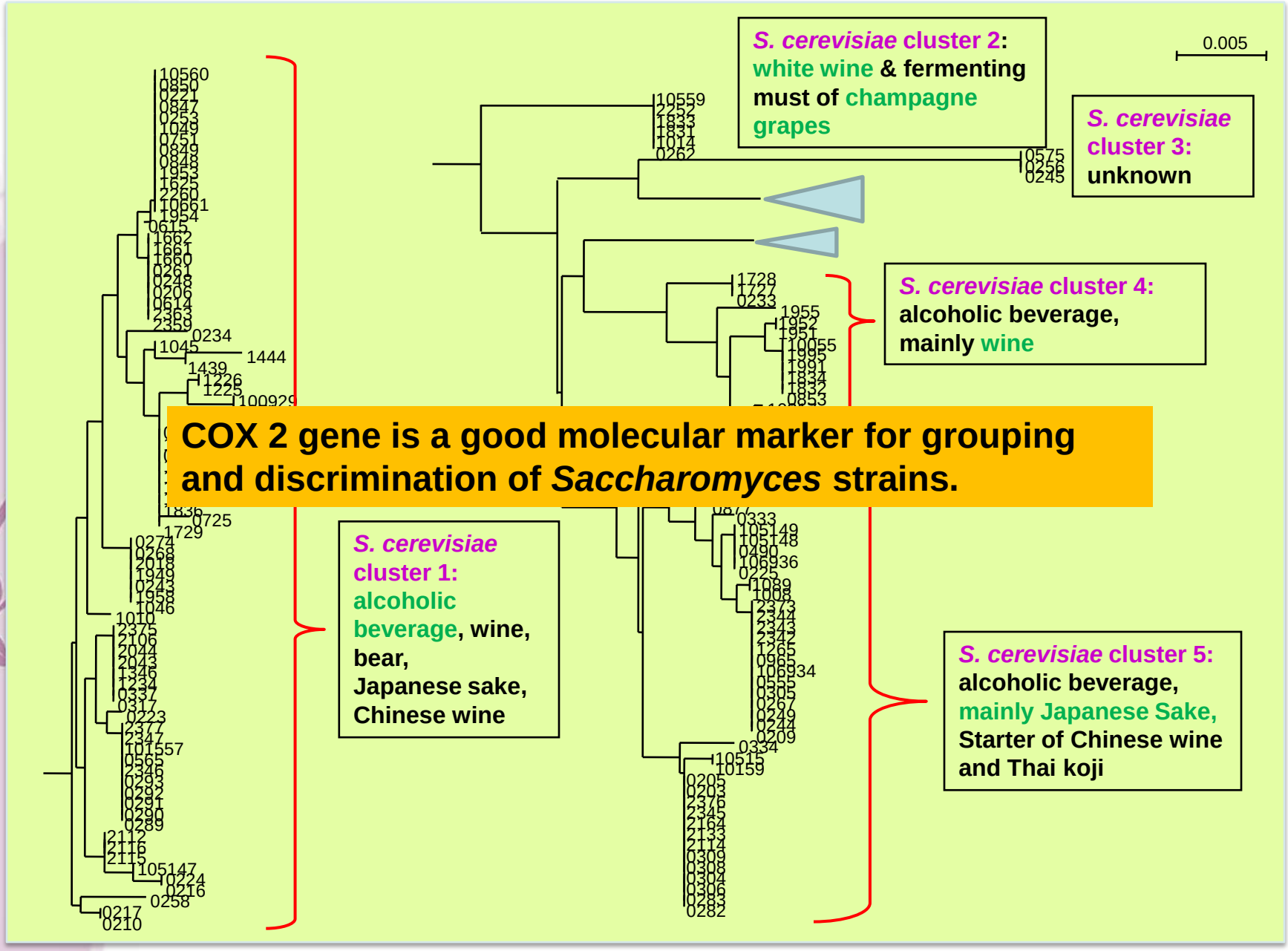
***Saccharomyces* species are monophyletic in the LSU r DNA tree.**

COX2 gene tree

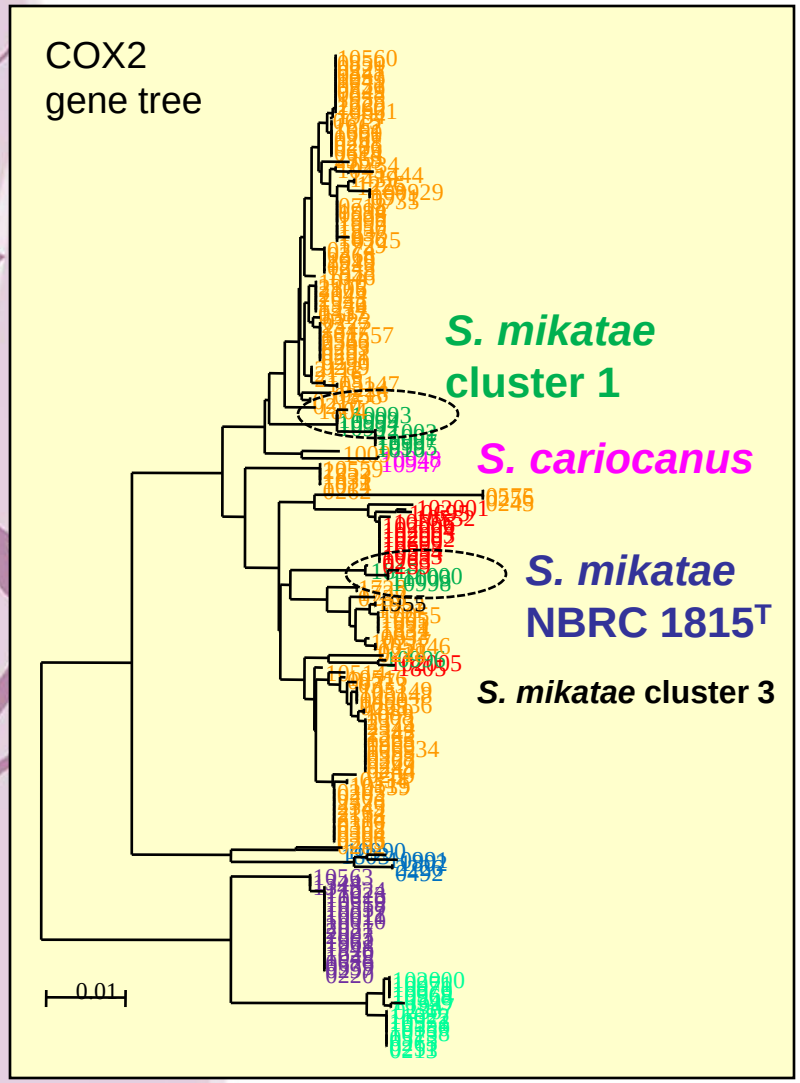
NJ /224 strains



Grouping within *S. cerevisiae* based on *Cox2*



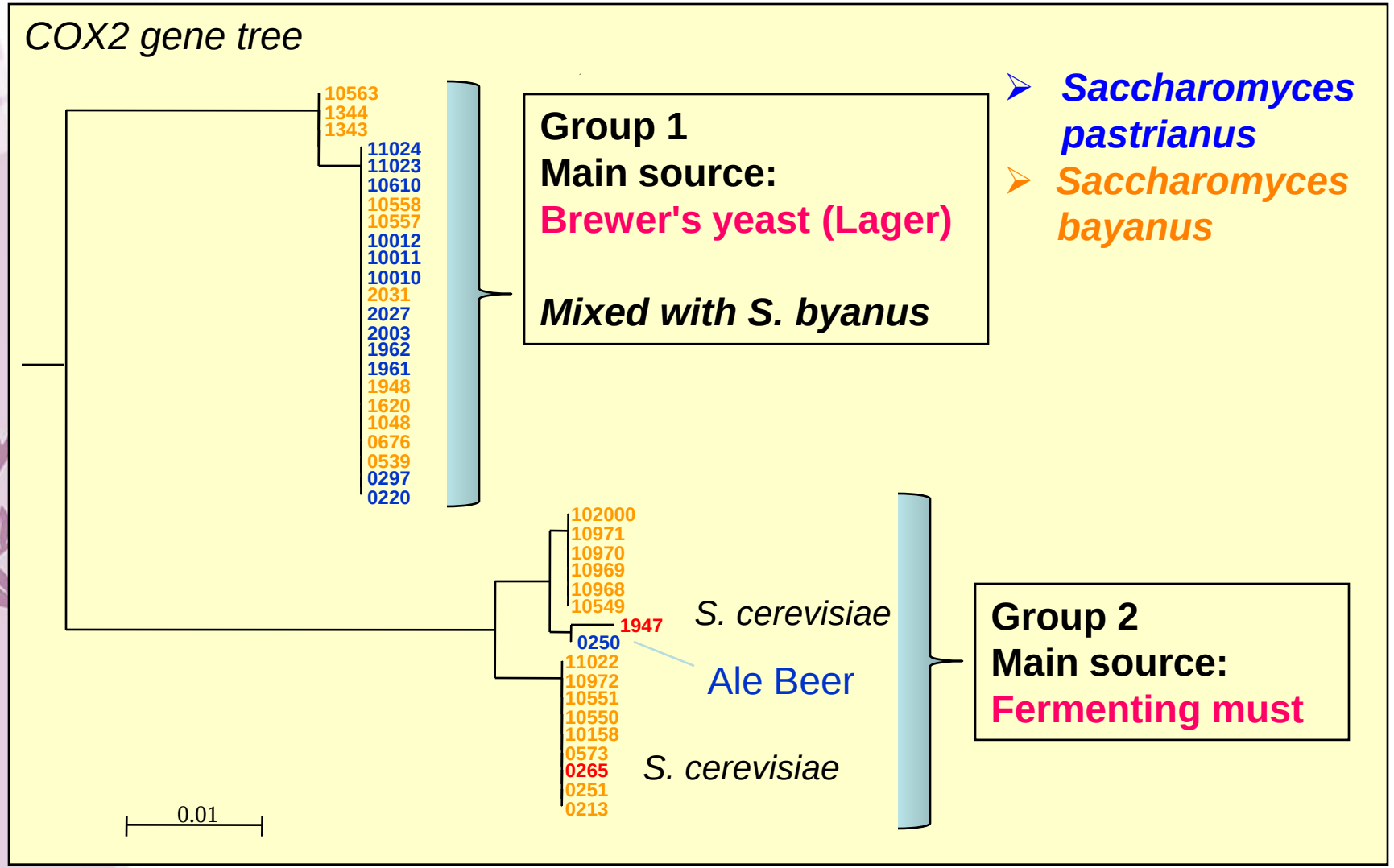
S. mikatae vs. *S. cariocanus*



	Source	Country
***S. mikatae* cluster 1**	Tree exudate Partially decayed leaf soil	Japan
***S. mikatae* cluster 2**	Soil Partially decayed leaf	Japan
***S. cariocanus* cluster 1**	Fruit fly	Brazil

	LSU rDNA D1/D2 (%)		ITS-5.8S-ITS2 (%)		*COX2* (%)		DNA-DNA hybrid. (%)	
	S. m.	***S. c.***	***S. m.***	***S. c.***	***S. m.***	***S. c.***	***S. m.***	***S. c.***
***S. mikatae* NBRC 1815T**	100	98	100	98	100	98	100	47
***S. mikatae* NBRC 10994**	98	98	99	98	97	97	84	48
***S. mikatae* NBRC 10999**	98	98	99	98	97	97	83	33
***S. cariocanus* NBRC 10947T**	98	100	98	100	98	100	32	100

Saccharomyces bayanus and *S. pastrianus*

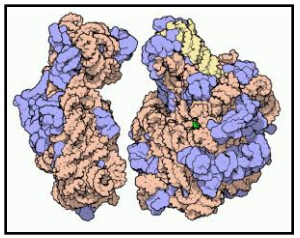


2. Ribosomal Subunit Protein Analysis using MALDI-TOF MS

- Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOFMS) is an emerging tool for microbe characterization and differentiation at the species and strain levels. Early reports focused on rapid identification of intact whole bacteria. Recently the MALDI-TOFMS analysis of intact yeast cells has been reported.
- On the other hand, variability of data by the difference of the culture condition is problem in the method of the above-mentioned.
- To solve this problem, we focus to the ribosomal subunit proteins as the biomarker for the identification of yeast.

Method

1. Wet cells (24 hours culture)
2. Purification of ribosomal protein
   using ultra centrifugation
3. Selection of bio-marker



Ribosomal protein

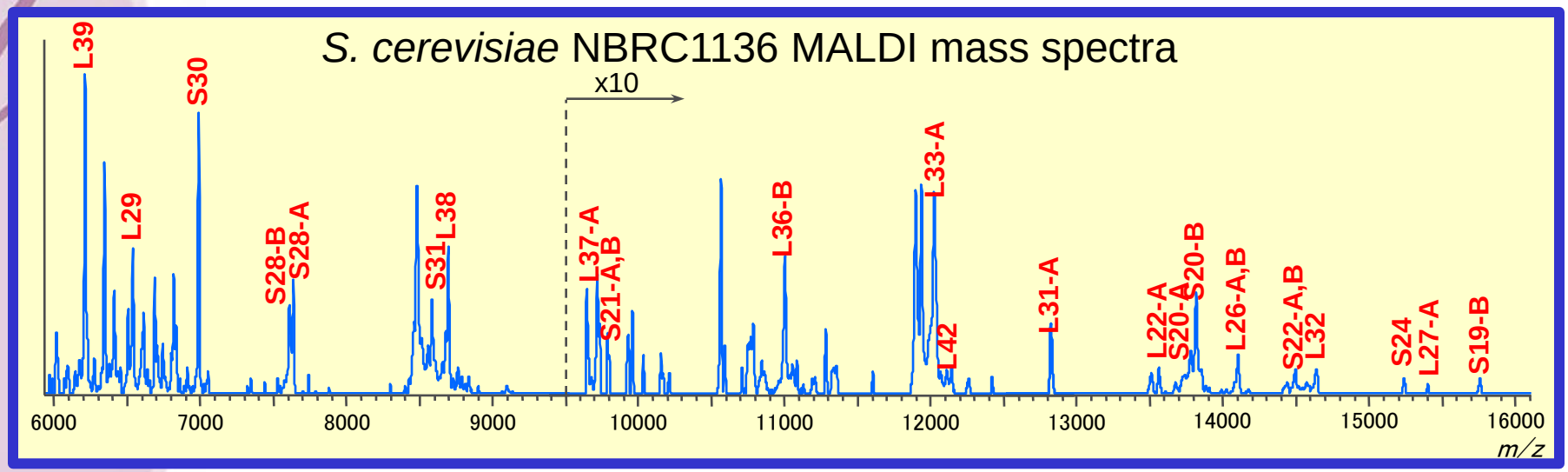
The theory value of the observation mass calculated based on the registration amino acid array of each ribosome subunit protein of genome sequenced *S. cerevisiae* NBRC 1136 was compared with the peak on the spectrum MALDI mass of the refinement ribosome protein. The one within 150ppm in error margin was used as biomarkers.

4. Mass spectral analysis of protein using MALDI-TOFMS

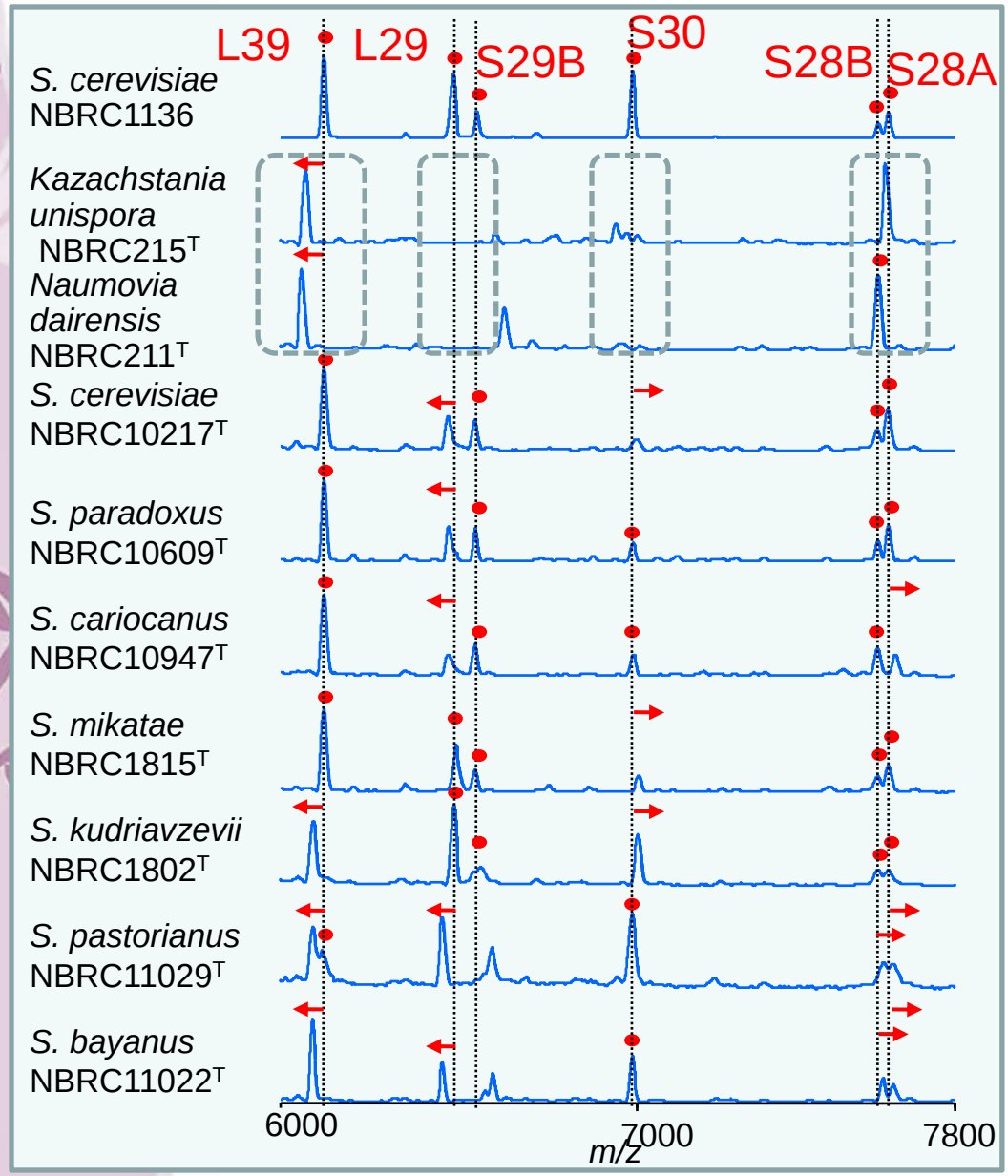
   matrix： sinapinic acid

   Range of measurement : *m/z* 6000-17000

(Teramoto *et al*., 2007)



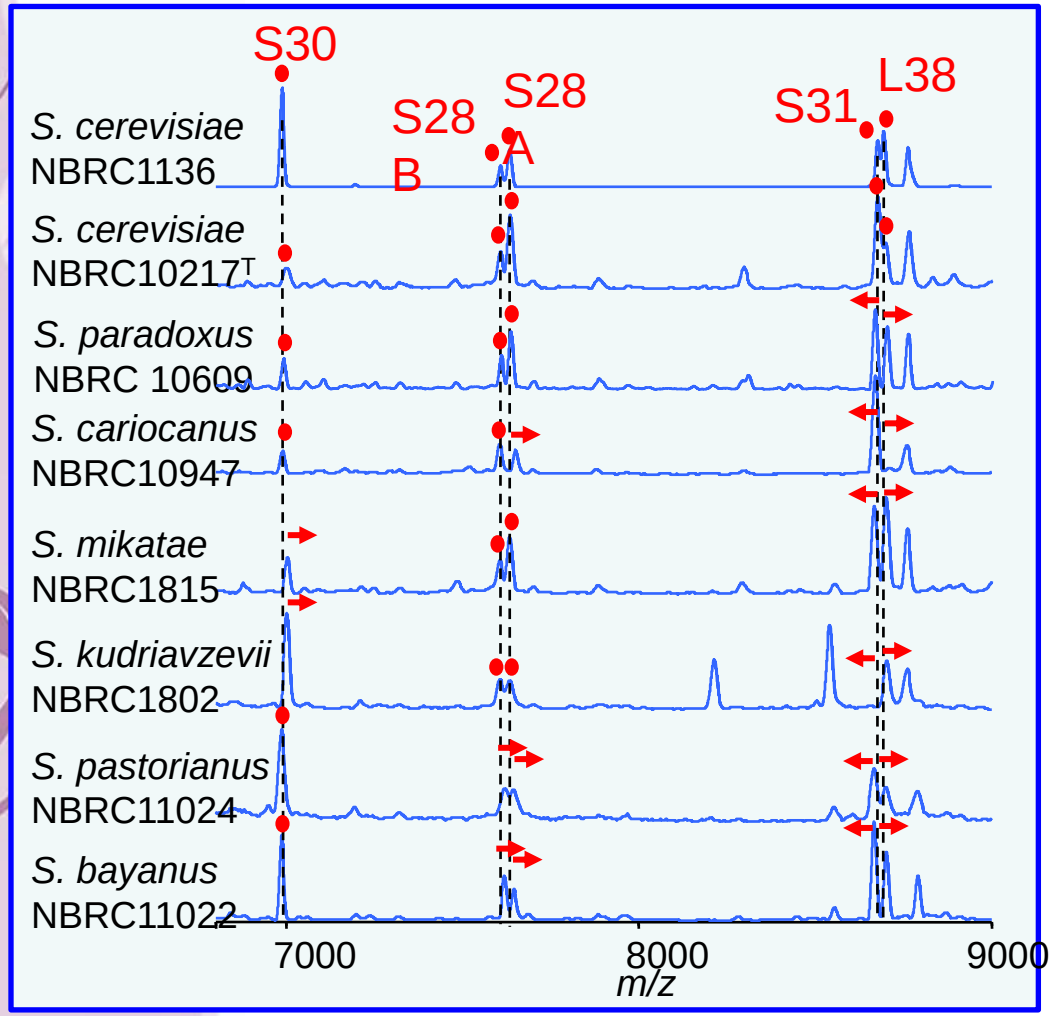
Result 1: Genus level

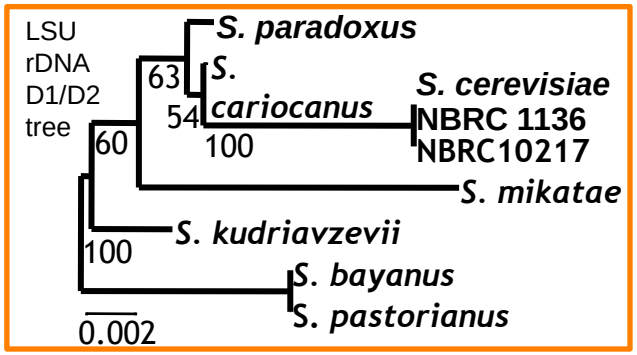


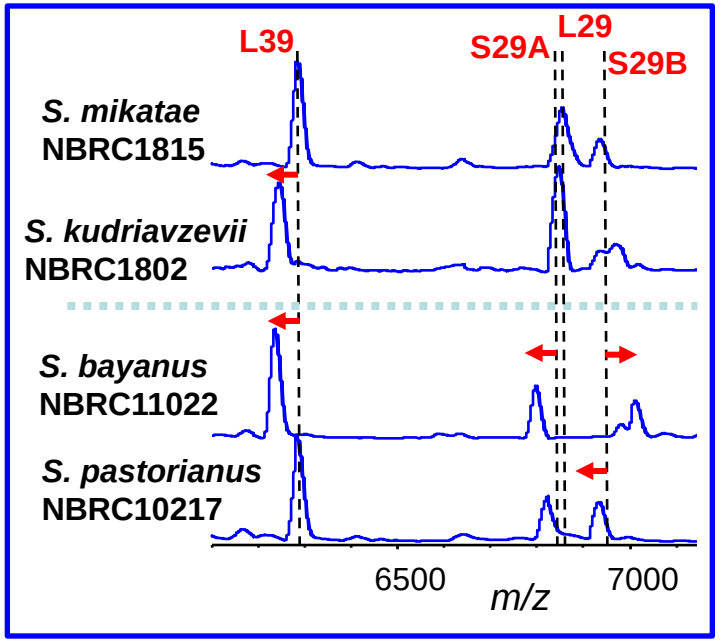
- Based on the *S. cerevisiae* whole genome sequence, 56 ribosome proteins (Small subunit) and 81 ribosome proteins (Large subunit) are presumed .
- In this study, the MALDI mass spectra were observed in the range of m/z, and used at least 40 kinds of proteins as molecular information.
- *Saccharomyces* species show clearly different molecular patterns from other related genera, *Naumovia* and *Kazachstania*.

Possibility to distinguish at genus level

Differentiation among *Saccharomyces* species

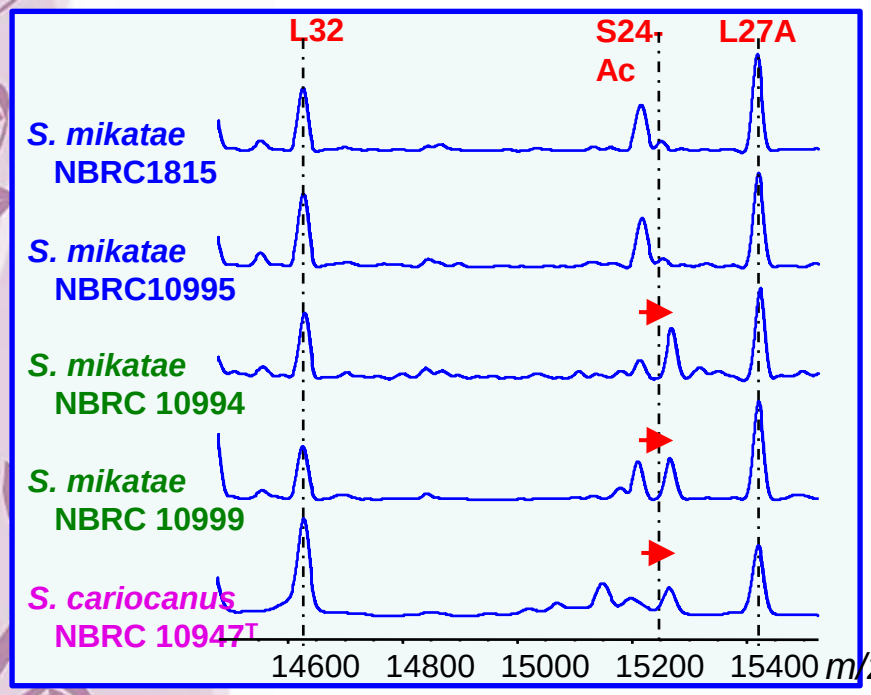


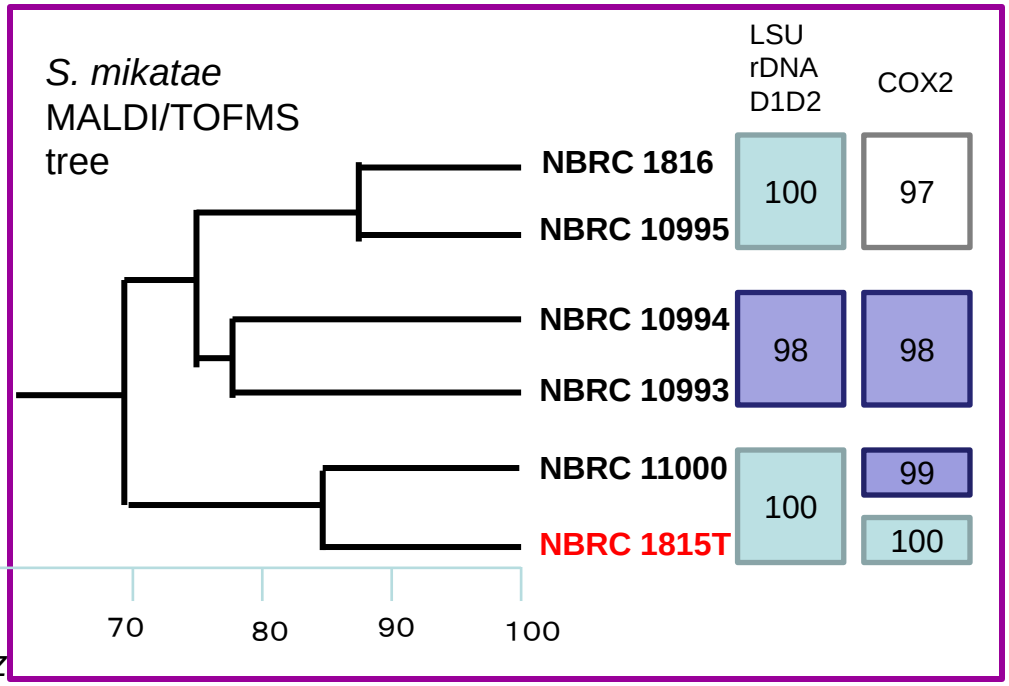




Six species of *Saccharomyces* show the different patterns in two or more ribosome subunit proteins. It shows the possibility to distinguish at species level.

Closely related strains: *S. mikatae* and *S. cariocanus*



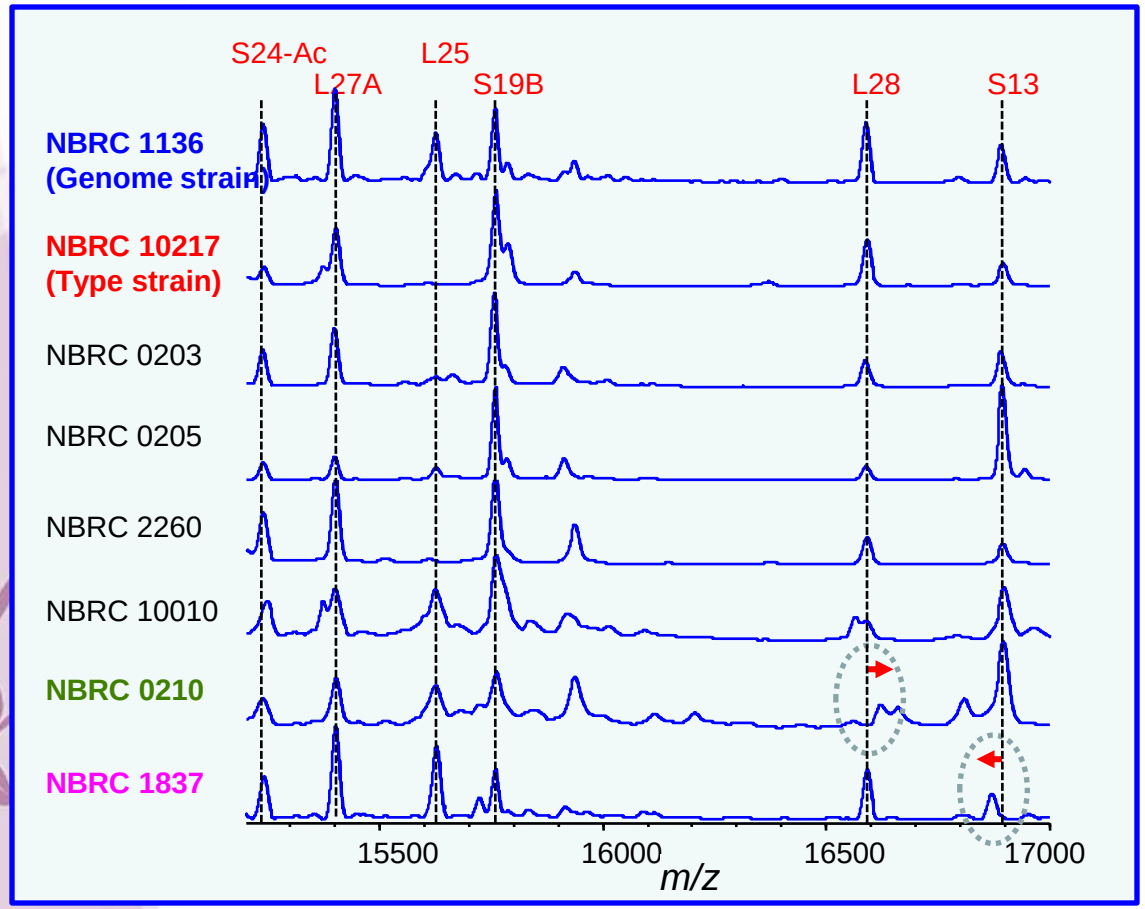


The *S. mikatae-S. cariocanus* immediate group (green) can be distinguished at S24 and S30, which result corresponded to the relationships based on the LSU rDNA D1/D2, ITS and COX2.

The MALDI/TOF MS result corresponded to the relationships based on the LSU rDNA D1/D2 and COX2.

⇨ **Possibility to distinguish at below the species level**

S. cerevisiae: at below the species level



NBRC no.	LSU rDNA D1/D2 Similarity
1136 (Genome strain)	-
10217	100%
0203, 0205, 2260, 10010	100%
0210	1base substitution
1837	2base substitution

All of the *S. cerevisiae* strains have similar MALDI mass spectra.
But, some strains which have some substitutions of LSU rDNA D1D2 (1-2 bases) show the different spectra at L28 and S13.

⇨ L28 or S13 will be a biomarker for grouping?

Conclusions

COX2 phylogenetic analysis and ribosomal subunit mass spectra analysis were applied to the grouping and discrimination of yeast *Saccharomyces*.

COX2:

The base substitution rate of the COX2 gene was larger than that of the LSU rDNA D1/D2. More than two clusters were formed in the homogeneous group. It show that COX2 becomes a molecular marker for grouping and discrimination. Now we have used the COX 2 gene analysis for the quality control of *Saccharomyces* group in NBRC yeast collection.

Ribosomal subunit proteins:

Several different peaks were observed in the *Saccharomyces* species used by MALDI-TOF MS. This method might be useful for the identification at the genus and species level. A few information can be used to distinguish in homogenous group.

It will be necessary to analyze the specific ribosomal subunit proteins for biomarker in each level. Future tasks are the promptness of the purification of the ribosome protein and making to the data base.