Nakazawaea atacamensis f.a., sp. nov. a novel non-conventional fermentative ascomycetous yeast species from the Atacama Desert

Francisco Cubillos[1], Macarena Araya[1], Tomás Moyano[2], Pablo Villarreal[1], Ana Raquel Santos[3], Francisca P Diaz[2], Andrea Bustos-Jarufe[2], Kamila Urbina[1], Marizeth Goenewald[4], Rodrigo A. Gutiérrez[2], and Carlos Rosa[3]

[1]Universidad de Santiago de Chile Facultad de Quimica y Biologia
[2]Instituto Milenio de Biologia Integrativa
[3]Universidade Federal de Minas Gerais Departamento de Microbiologia
[4]Westerdijk Fungal Biodiversity Institute

October 7, 2023

Abstract

Atacama is the most hyper-arid Desert in the world. In this study, we describe a novel species, *Nakazawaea atacamensis* f. a., sp. nov., isolated from plant samples in the Atacama Desert of Chile. In total, three isolates of *N. atacamensis* were obtained from independent *Neltuma chilensis* bark samples (synonym Prosopis chilensis, Algarrobo). The novel species was delineated based on morphological, physiological, biochemical, and molecular characteristics. A neighbour-joining analysis using the sequences of the D1/D2 domains of LSU rRNA revealed that *N. atacamensis* sp. nov. clustered with *Nakazawaea pomicola*. The sequence of *N. atacamensis* differed from closely related species by 1.3% to 5.2% in the D1/D2 domains. A phylogenomic analysis based on single nucleotide polymorphism's data confirms that the novel species belongs to the genus *Nakazawaea,* and placed *N. atacamensis* closer to *N. peltata*. Phenotypic comparisons demonstrated that *N. atacamensis* sp. nov. exhibited distinct carbon assimilation patterns compared to its related species. Genome sequencing of the ATA-11A-B ^{T} strain revealed a genome size of approximately 12.4 Mbp, similar to other *Nakazawaea* species, with 5,116 protein-coding genes annotated using InterProScan. In addition, *N. atacamensis* exhibited the capacity to ferment synthetic wine must, representing a potential new yeast for mono or co-culture wine fermentations. This comprehensive study expands our understanding of the genus *Nakazawaea* and highlights the ecological and industrial potential of these yeasts in fermentation processes. The holotype of *N. atacamensis* sp. nov. is CBS 18375 ^{T}. The Mycobank number is MB 849680.

Nakazawaea atacamensis f.a., sp. nov. a novel non-conventional fermentative ascomycetous yeast species from the Atacama Desert

Macarena Araya [1], Tomás Moyano [2,3], Pablo Villarreal [1,2], Ana R. O. Santos[4], Francisca P. Díaz [2,5,6], Andrea Bustos-Jarufe [2], Kamila Urbina [1,7], Marizeth Groenewald [8], Rodrigo A. Gutiérrez[2,3,9], Carlos A. Rosa [4], Francisco A. Cubillos [1,2,7+]

[1]*Universidad de Santiago de Chile, Facultad de Quimica y Biologia, Departamento de Biologia, Santiago, Chile.*

[2]*Millennium Institute for Integrative Biology (iBio), Santiago, Chile*

[3]*Departamento de Genetica Molecular y Microbiologia, Facultad de Ciencias Biologicas, Pontificia Universidad Catolica de Chile, Santiago 7500565, Chile*

[4]*Departamento de Microbiologia, ICB, C.P. 486, Universidade Federal de Minas Gerais, Belo Horizonte, MG, 31270-901, Brazil*

[5] *Instituto de Geografia, Pontificia Universidad Catolica de Valparaiso, Valparaiso, Chile.* [6]*Nucleo Milenio de Ecologia Historica Aplicada para los Bosques Aridos (AFOREST), Santiago, Chile*

[7]*Millennium Nucleus of Patagonian Limit of Life (LiLi), Valdivia, Chile*

[8]*Westerdijk Fungal Biodiversity Institute, 3584 Utrecht, The Netherlands*

[9]*Millennium Institute Center for Genome Regulation. Institute of Ecology and Biodiversity*

+ Corresponding author. Email address :francisco.cubillos.r@usach.cl

Running title: Novel species *Nakazawaea atacamensis*

Keywords: yeast, Atacama, fermentation, novel species.

TAKE AWAY

- A novel species, *Nakazawaea atacamensis* sp. nov., was isolated from *Neltuma chilensis* (Algarrobo) samples in the Atacama Desert of Chile
- The sequence of *N. atacamensis* differs from closely related species by 1.3% to 5.2% in the D1/D2 domains .
- *N. atacamensis* exhibited the capacity to ferment synthetic wine must, suggesting its potential for wine fermentations.

ABSTRACT

Atacama is the most hyper-arid Desert in the world. In this study, we describe a novel species, *Nakazawaea atacamensis* f. a., sp. nov., isolated from plant samples in the Atacama Desert of Chile. In total, three isolates of *N. atacamensis* were obtained from independent*Neltuma chilensis* bark samples (synonym Prosopis chilensis, Algarrobo). The novel species was delineated based on morphological, physiological, biochemical, and molecular characteristics. A neighbour-joining analysis using the sequences of the D1/D2 domains of LSU rRNA revealed that *N. atacamensis* sp. nov. clustered with*Nakazawaea pomicola* . The sequence of *N. atacamensis*differed from closely related species by 1.3% to 5.2% in the D1/D2 domains. A phylogenomic analysis based on single nucleotide polymorphism's data confirms that the novel species belongs to the genus*Nakazawaea,* and placed *N. atacamensis* closer to *N. peltata* . Phenotypic comparisons demonstrated that *N. atacamensis*sp. nov. exhibited distinct carbon assimilation patterns compared to its related species. Genome sequencing of the ATA-11A-B^{T}strain revealed a genome size of approximately 12.4 Mbp, similar to other *Nakazawaea* species, with 5,116 protein-coding genes annotated using InterProScan. In addition, *N. atacamensis*exhibited the capacity to ferment synthetic wine must, representing a potential new yeast for mono or co-culture wine fermentations. This comprehensive study expands our understanding of the genus*Nakazawaea* and highlights the ecological and industrial potential of these yeasts in fermentation processes. The holotype of *N. atacamensis* sp. nov. is CBS 18375^{T}. The Mycobank number is MB 849680.

INTRODUCTION

Yeasts are known for their ubiquitous presence in various biomes, where they inhabit a wide range of substrates and environments. These include tropical forests (Morais *et al.* , 2006), temperate forests (Mozzachiodi *et al.* , 2022), hyperarid deserts, and cold habitats (Buzzini *et al.* , 2018). Hyperarid desert habitats are characterized by challenging conditions such as low oxygen and water levels, extreme temperatures, and exposure to ultraviolet (UV) radiation (Houston & Hartley, 2003, Schulze-Makuch *et al.* , 2018). To survive in these conditions, yeasts have evolved various adaptation mechanisms. For instance, they produce carotenoid pigments and mycosporines to withstand UV radiation, and possess aquaporins to counteract sudden osmotic changes (Aponte-Santamaria *et al.* , 2017). However, despite the importance of yeast diversity and their metabolic capacity to tolerate environmental stresses and nutrient limitations, most studies in hyperarid deserts have focused on bacteria, neglecting the role of yeasts, especially in extreme environments such as deserts (Drees *et al.* , 2006, Finstad *et al.* , 2017, Azua-Bustos *et al.* , 2018).

Exploring microbial diversity in inhospitable habitats represents a promising strategy for discovering novel yeast species with valuable biotechnological potential. In this regard, ethanol-tolerant yeasts demonstrated remarkable abilities that revolutionized biotechnology industries, for example by providing biological and genetic resources for the production of novel beverages with unique organoleptic properties (Libkind *et al.* , 2011, Cubillos *et al.* , 2019). For instance, a recent bioprospecting study conducted in the Patagonian rainforest led to the identification of new *Saccharomyces*species, including *S. eubayanus* , which has proven valuable for the brewing industry (Libkind *et al.* , 2011, Peris *et al.* , 2016, Eizaguirre *et al.* , 2018, Nespolo *et al.* , 2020). Similarly, the recent advent of non-*Saccharomyces* species from natural environments is becoming an interesting biological resource to develop novel fermented beverages (Canonico *et al.* , 2019, Villarreal *et al.* , 2022). However, the occurrence and diversity of culturable yeasts in other environments, particularly hyper-arid and desert areas, have been largely overlooked (Wei *et al.* , 2022).

The Atacama Desert, situated in northern Chile (19o-27oS), stands as the driest and oldest desert on Earth, representing an extreme environment for life (Ewing *et al.* , 2006). The central valley is classified as a hyper-arid desert with an average annual precipitation of less than 20 mm. The Atacama Desert experiences daily temperature fluctuations ranging from -6 degC to 38 degC (McKay *et al.* , 2003). Moreover, it is characterized by intense UV radiation, limited organic carbon sources, and low water availability (Navarro-Gonzalez *et al.* , 2003). Despite these harsh conditions, a few resilient woody plants and perennial grasses manage to thrive in this environment (Carrasco-Puga*et al.* , 2021). Similarly, the Atacama Desert soil harbours diverse microorganisms, particularly in surficial salt crusts (Wierzchos*et al.* , 2011) and in association with plants (Eshel *et al.* , 2021). Previous studies have predominantly focused on bacteria in the field of Astrobiology, aiming to gain fundamental insights into terrestrial desert environments and their microbiota (Azua-Bustos*et al.* , 2012, Pulschen *et al.* , 2015). However, investigations on yeast communities associated with the flora of the Atacama Desert, thriving under extremely arid conditions, remain scarce.

Historically, the indigenous people of the Atacama region have used various plants to produce spontaneously fermented beverages, which typically exhibit low ethanol levels. These beverages are created through the natural fermentation of leguminous tree pods and fruit juices, providing a source of calories, proteins, vitamins, minerals, and diverse bioactive compounds (Sciammaro *et al.* , 2016). For instance, 'Aloja' is an alcoholic beverage made by fermenting mashed pods of the *Neltuma* (synonym *Prosopis*) tree, locally known as Algarrobo (Pardo O., 2015, Sciammaro *et al.* , 2022). This practice suggests the presence of indigenous ethanol-tolerant yeasts within naturally fermented musts, indicating a promising opportunity for the discovery of novel fermentative yeasts within the local vegetation of the Atacama Desert.

In this study, our objective was to isolate ethanol-tolerant yeasts associated with the flora of the Atacama Desert using an enrichment culture strategy. We successfully obtained three yeasts isolates from different tree samples, one of them from bark-exuded gum and two from barks obtained from independent *N. chilensis* trees. Further investigations revealed that these isolates represent a previously unidentified yeast species. Morphological and physiological characteristics, as well as phylogenetic analyses, provide strong support for classifying these three strains as a new ascomycetous yeast species within the genus *Nakazawaea* . We propose the name*Nakazawaea atacamensis* sp. nov. for this novel species. Furthermore, we demonstrate the potential fermentation utilization of this species.

2. MATERIALS AND METHODS

Sampling and yeast isolation

The sampling for this study took place in the Atacama Desert near the San Pedro de Atacama village (22deg 55' S, 68deg 12' W) in January 2021. The sampling strategy involved identifying tree species and collecting various samples from each tree. Briefly, 5 g of bark, bark-exuded gum, pods, and flowers of Algarrobo (*Neltuma chilensis* (Molina) C.E.Hughes & G.P.Lewis), Tamarugo (*Strombocarpa tamarugo* (Phil) C.E.Hughes & G.P.Lewis) and Chanar (*Geoffroea decorticans* were collected. A total of 22 samples were collected in aseptic conditions and transferred to tubes containing 10 mL of selective enrichment medium composed of YNB (Yeast Nitrogen Base) supplemented with 2% (w/v) glucose and 4% (v/v) ethanol. This

medium was specifically designed to exclusively select for ethanol-tolerant yeast species (Villarreal *et al.* , 2022). The tubes were incubated at 25degC without agitation for a period of 14 days. After incubation, 100 µl aliquots were spread onto yeast-extract peptone dextrose (YPD) agar plates containing chloramphenicol (20 µg ml-1) and incubated at 25°C until yeast colonies emerged. Representative colonies of each distinct morphotype were purified through streak inoculation on YPD agar for further characterization. To ensure long-term preservation, yeast cultures were stored at -80°C in a broth culture supplemented with 20% (w/v) glycerol.

Sanger sequencing and phenotypic characterization

Part of the small subunit (SSU) rRNA gene, the internal transcribed spacer (ITS) region and the D1/D2 domains of the large subunit (LSU) rRNA genes were amplified and sequenced using primers NS1 and NS4 (White*et al.* , 1990, Kurtzman & Robnett, 1998, Lachance *et al.* , 1999), ITS1 and ITS4 (White *et al.* , 1990) and LR0R and LR16 (Vilgalys & Hester, 1990, Moncalvo *et al.* , 2000), respectively. The PCR products were purified using QIAquick PCR columns (Qiagen) following the manufacturer's instructions. After purification, the products were sent to Macrogen (Korea) for sequencing, employing the respective PCR primers. To identify the yeast species, the sequences of the SSU, ITS region, and the D1/D2 domains of the LSU rRNA gene were compared with those available in GenBank. This comparative analysis was conducted using the 'blastn' search utility (McGinnis & Madden, 2004). To identify novel species, we used a cut-off of 97% sequence identity (STACKEBRANDT & GOEBEL, 1994, Vu *et al.* , 2016, Lachance, 2018). All sequences generated during the study were deposited in NCBI GenBank. The GenBank accession numbers of the ITS, SSU, and LSU rDNA sequences are OP293325, OP293328, and OP293331 for ATA-11A-B (=CBS 18375^{T}), OP293326, OP293329, and OP293332 for isolated ATA-12C-B (=CBS 18376) and OP293327, OP293330, and OP293333 for the ATA-13E-S (=CBS 18374) strain, respectively.

For phylogenetic analysis, only *Nakazawaea* species that exhibited sequences of the SSU rRNA gene, the ITS region, and the D1/D2 domains of the LSU rRNA gene were included. Sequences were edited, assembled, concatenated, and aligned using the MUSCLE multiple alignment program in MEGA software version 11 (Kumar *et al.* , 2018). The phylogenetic relationship of the novel species was determined through Neighbor-Joining analysis, based on the D1/D2 domains of the LSU rRNA gene, and using *Pachysolen tannophilus* as the outgroup species. For this analysis, the number of substitutions between the sequences was used as the distance metric. Confidence values were estimated from bootstrap analyses of 1,000 replicates (Felsenstein, 1985). Nodes were considered supported if the bootstrap percentage was [?]50 % (Hillis & Bull, 1993). The identity matrix between *Nakazawaea* species was generated with Bioedit (Hall *et al.* , 2011).

The yeasts were subjected to morphological, physiological, and biochemical characterization under solid media conditions using the standardized methods outlined by Kurtzman et al. (2011). To assess their fermentative capacity, the ability to metabolize glucose, fructose, and sucrose was examined in Durham tubes containing fermentation base media, with a final sugar concentration of 2% (w/v), as described by Yarrow (1998). The tubes were incubated at 25degC for a period of 14 days. For cell morphology analysis, observations were made using a Nikon Eclipse Ti2-E microscope equipped with differential interference contrast (DIC) optics after 3 days of growth in YPD broth, incubated at 25degC.

Whole genome sequencing

For genome sequencing, the strain ATA-11A-B (=CBS 18375^{T}) was cultured in 15 ml of YPD broth at 25degC for 72 hours. Genomic DNA extraction was performed using the Qiagen Genomic-tip 20/G kit (Qiagen, Germany), following the previously described method by Nespolo et al. in 2020 (Nespolo *et al.* , 2020). The extracted DNA was quantified using the Invitrogen Qubit 4 Fluorometer (cat. no. Q33226), and a genomic DNA library was prepared using the Illumina DNA-prep kit, following the manufacturer's protocol. Subsequently, the libraries were sequenced using the Illumina NextSeq500 with a Mid-Output Kit, generating 150-bp paired-end reads, at the sequencing facility of the University of Santiago de Chile.

Genome assembly, gene function annotation and phylogenetic analyses

A total of 7,225,123 raw Illumina sequences were obtained, filtered, and cleaned using the trim_galore program v0.6.7 (https://github.com/FelixKrueger/TrimGalore). Out of these, 7,206,640 sequences, with both pairs preserved, were utilized for genome assembly. The de novo genome assembly was conducted using the multi-platform genome assembly pipeline (MpGAP) v3.1, employing Nextflow version 21.10.6 and Masurca version 4.0.5. To assess the quality of the assembly, Quast version 5.0.2 and Busco version 5.4.2 were employed, utilizing the Saccharomycetales_odb10 database. The assembly with the best performance (highest N50 and largest contig size (Kbp)) was then annotated using the Funannotate pipeline version 1.8.14. Contigs smaller than 500 base pairs were removed, and repeated sequences were masked using default settings. Gene prediction was carried out ab initio using 'Augustus', 'HiQ', 'GlimmerHMM', 'snap', and 'GeneMark'. The prediction of t-RNAs was performed using tRNAscan-SE v2.0.9, a program included in Funannotate. For annotation and functional prediction of genes we utilized InterProScan. KEGG and KofamKOALA web server were used to predict the gene functions of *N. atacamensis* ATA-11A-B. The average nucleotide identity (ANI) between *Nakazawaea* genomes was estimated from different available assemblies using OrthoANI (Lee*et al.* , 2016).

A maximum-likelihood phylogenetic tree (ML) was constructed using a protein sequence predicted from *N. atacamensis* and other four*Nakazawaea* species (*N. peltata* , *N. holstii* ,*N. ishiwadae* , and *N. ambrosiae*). The yeast species*P. tannophilus* was used as an outgroup. Ortho-Finder v2.4.1 was employed to identify orthologous protein groups among these six different species. Subsequently, a total of 2,422 single-copy orthologs were identified in all species and aligned using Muscle v3.8.15 (Edgar, 2004). Alignments were concatenated to produce a maximum-likelihood tree with RAxML v8.2.12 (-f a -x 12345 -p 12345 -# 100 -m PROTGAMMAJTT -k). The phylogenetic tree was visualized and plotted using iTOL v5.

Microculture and fermentative phenotypic characterization

The microculture assay was performed in liquid media as previously described (Nespolo *et al.* , 2020). Briefly, isolates were pre-cultivated in 200 µl 0.67% YNB medium supplemented with glucose 2% for 48 h at 25°C. Each pre-inoculum (optical density (OD) of 0.03–0.1) from *N. atacamensis* isolates was inoculated in 200 µL of media with the following carbon sources: glucose 2%, fructose 2% and sucrose 2% for 64 h incubated without agitation using a Tecan Sunrise absorbance microplate reader. Additionally, we included environmental stressors such as ethanol 4%, 6% and 8%, and glucose 20% during 64 h. The OD was measured every 30 minutes using a 630 nm filter. Each experiment was carried out in triplicate. Maximum growth rate, lag time, and OD max parameters were obtained for each isolate using the GrowthRates software as previously described (Villarreal *et al.* , 2022). All statistical analyses were performed using biological replicates. One-way ANOVAs (Analysis of Variance) were performed using GraphPad Prism 8.01 for Windows (GraphPad Software, La Jolla, CA, USA,*www.graphpad.com*).

Fermentations were carried out as previously described (Villarreal*et al.* , 2022). Briefly, for each experiment, yeast cells were initially grown under constant agitation in 10 mL of SWM for 16 hours at 25°C. Next, 1x10^6 cells/mL were inoculated into 50 mL SWM (in 250 mL flasks) and incubated at 25°C with constant agitation for 7 days. As wine fermentation control, we used the strain *S. cerevisiae* EC1118 strain. The experiments were carried out in triplicates. Micro-fermentations were weighed every day to calculate the CO_2 output. Sugar consumption and metabolite production were estimated using HPLC and a Bio-Rad HPX-87H column. In this way, we estimate the consumption of glucose and fructose, together with the production of glycerol and ethanol.

3. RESULTS AND DISCUSSION

3.1 Isolation of ethanol-tolerant yeast species from Atacama Desert flora

To identify ethanol-tolerant yeast strains from the Atacama Desert, we collected 22 samples from various tissues of native trees, including bark, bark-exuded gum, pods, and flowers. The tree species sampled included *Neltuma chilensis* (Algarrobo), *Geoffroea decorticans* (Chañar) and *Strombocarpa tamarugo* (Tamarugo). After subjecting these samples to culture enrichment in rich media containing 4% ethanol, a total of 91 ethanol-tolerant yeast colonies were obtained (**Table S1**). Among these colonies, 60 were isolated from*N.*

chilensis , 10 from *S. tamarugo* , and 21 from *G. decorticans* .

Next, we amplified the ITS region and sequenced at least one colony per sample according to the size of the ITS region obtained. Based on these sequencing results, we identified 36 isolates as belonging to the basidiomycetous phylum, specifically the genera *Kwoniella* ,*Rhodotorula,* and *Naganishia.* Additionally, we identified 55 isolates corresponding to ascomycetous species, belonging to the genera *Zalaria, Starmerella* , and *Nakazawaea* .

To identify fermenting yeast, we conducted the Durham tube fermentation test using glucose as the carbon source. Through this test, we discovered thirty-nine isolates belonging to the genera*Nakazawaea* and *Starmerella* that exhibited the ability to ferment glucose. Notably, all of these isolates were obtained from samples of *N. chilensis* . This plant is a leguminous tree known for its high levels of polysaccharides (Astudillo *et al.* , 2000). It is also worth mentioning that the 'Aloja' beverage is produced from the mashed pods of *Neltuma* species, suggesting a potential association between ethanol-tolerant yeasts and the spontaneous fermentation processes associated with this tree species. Among the isolates, the majority were identified as representatives of the genus*Nakazawaea* (35 isolates). However, the ITS region of these isolates differed by 16 to 25% in terms of sequence identity from the*Nakazawaea* species currently described (**Figure S1**). Hence, in this study, we focus on the description of a new fermenting ascomycetous yeast species, which we have named *Nakazawaea atacamensis* sp. nov.

3.2. Novel species delineation and Identification

Delineation of a new species and phylogenetic placement of *Nakazawaea atacamensis* sp. nov.

Three *Nakazawaea* isolates from different tree samples, representing a novel anamorphic and sucrose-fermenting yeast species, were isolated from *N. chilensis* samples in the Atacama Desert of Chile. Two isolates of *N. atacamensis* were obtained from bark samples from two different trees (ATA-11A-B^{T} and ATA-12C-B), while the third isolate was obtained from bark-exuded gum from another tree (ATA-13E-S). To assess the sequence divergence among the three *N. atacamensis* isolates, we compared the ITS region, the small subunit (SSU) rRNA gene sequence, and the D1/D2 domains of the LSU rRNA gene. We found that the sequences of the internal transcribed spacer (ITS) and the LSU rRNA domains were 100% identical among the three strains (**Figure S2**). However, in the SSU rRNA region, the ATA-11A-B^{T} and ATA-12C-B isolates, both derived from bark samples, exhibited a 99.8% sequence identity compared to ATA-13E-S isolate (**Figure S3**). Strain CBS 5808 (GenBank accession number AY366526) and the novel species have identical D1/D2 sequences. This strain was deposited in 1970 by J. Grinbergs in the CBS Yeast Collection as *Candida conglobata* , and it was isolated from tree bark in Chile. The ITS sequences of strain CBS 5808 and the novel species were also identical, showing that they represent the same yeast species.

To determine the presence of the new *Nakazawaea* species in a phylogenetic context, we conducted a comparative analysis using the D1/D2 domains of the LSU rRNA gene from the currently accepted*Nakazawaea* species. The obtained sequences were utilized to construct a Neighbor-Joining tree. This tree highlighted the clustering of *N. atacamensis* with *N. pomicola* , and also indicated that *N. atacamensis* represents a novel species (**Figure 1**). The purpose of the Neighbor-Joining tree shown in Figure 1 is to demonstrate that our *Nakazawaea* isolates are indicative of a new species rather than to establish a comprehensive phylogeny of the genus. By comparing the DNA sequences of the investigated *N. atacamensis*strains with those of the genus, notable differences were observed. Specifically, *N. atacamensis* strains exhibited sequence divergences to the other species in the genus ranging from 1.3% to 5.2% in the D1/D2 domains. These findings provide strong evidence that the investigated strains represent a distinct anamorphic species within the *Nakazawaea* clade. Therefore, we propose the name*Nakazawaea atacamensis* f. a., sp. nov. for these three isolates. The mention *forma asexualis* (f. a.) is added as a reminder that the sexual state is not known (Lachance, 2012).

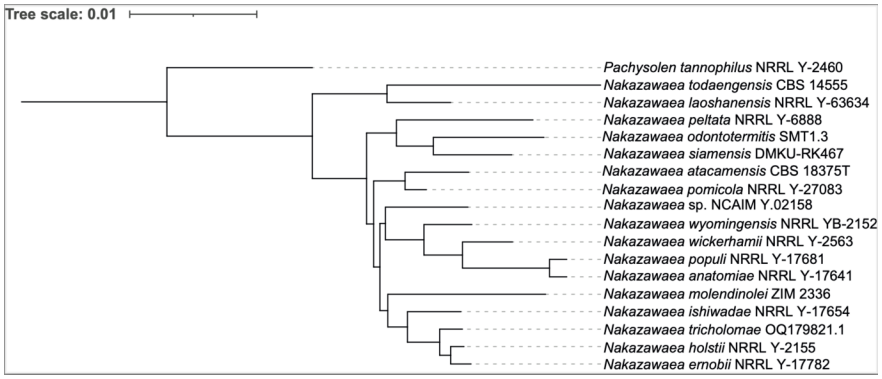


Figure 1. Neighbour-joining phylogram showing the placement of*Nakazawaea atacamensis* sp. nov sp. nov. within the genus . The tree was built based on the sequences of the D1/D2 domains of the LSU rRNA gene. The alignment was performed with MUSCLE. *Pachysolen tannophilus* was used as an outgroup. The distance metric is the number of substitutions. Bar, Substitutions per site.

3.3. Taxonomy

3.3.1. *N akazawaea atacamensis* M. Araya, T. Moyano, P. Villareal, A.R.O. Santos, FP. Díaz, A. Bustos-Jarufe, K. Urbina, R.A. Gutiérrez, C.A. Rosa and F.A. Cubillos sp. nov.

Etymology: *N akazawaea atacamensis*

(a.ta.ca.men'sis, NL. fem. Adj. *atacamensis*, pertaining to the Atacama Desert biome)

Isolates ATA-11A-B, ATA-12C-B, and ATA-13E-S displayed morphological characteristics consistent with those commonly observed in species of the genus *Nakazawaea.* After 3 days on YM agar at 25 ºC colonies are small, convex, grayish-white, and have an entire margin. Cells are ellipsoidal to elongate (1.1 - 3.1 x 2.1 – 4.5 µm) and occur singly or in mother-bud pairs (**Figure 2**). Budding is multilateral, though predominantly polar. In Dalmau plates after 14 days on corn meal agar, pseudohyphae and true hyphae were not observed. No ascospores or signs of conjugation were not seen on sporulation media. Fermentation of glucose, fructose and galactose is positive, while maltose, trehalose and sucrose are negative (**Table S2**). The species utilizes D-glucose, sucrose, D-galactose, α-trehalose, maltose, melezitose, cellobiose, salicin, L-rhamnose, D-xylose, L-arabinose, D-arabinose (w/s), D-ribose, ethanol, glycerol, erythritol (slow), ribitol, D-mannitol, D-glucitol, succinate, citrate, D-gluconate (slow), xylitol (slow), ethyl acetate (w/s) and N-acetyl-D-glucosamine as carbon sources. No growth occurs on inulin, raffinose, melibiose, lactose, soluble starch, L-sorbose, methanol, galactitol, myo-inositol, DL-lactate, D-glucosamine, hexadecane, acetone and 2-propanol. Lysine is utilized as sole nitrogen source but not nitrate and nitrite. Growth in amino-acid-free medium is positive. Growth at 10, 20, and 35°C is positive. Growth on YPD agar with 10 % sodium chloride and growth on 50 % glucose/yeast extract (0.5 %) are negative. Acid production is positive (slow). Starch-like compounds are absent. Growth in the presence of 0.01% cycloheximide is positive. Diazonium blue B color and urease reactions are negative. The holotype of *Nakazawaea atacamensis* , strain CBS 18375^{T}, is preserved in a metabolically inactive state at the CBS Yeast Collection of the Westerdijk Fungal Diversity Institute in Utrecht, the Netherlands. Additionally, an isotype of *N. atacamensis* has been deposited as strain RGM 3383 in the Chilean Culture Collection of Microbial Genetic Resources (CChRGM) at the Agricultural Research Institute (INIA) in Chile. The Mycobank number is MB 849680. Furthermore, 35 paratypes of*N. atacamensis* were isolated and are maintained in a metabolically inactive state at the Universidad de Santiago de Chile.

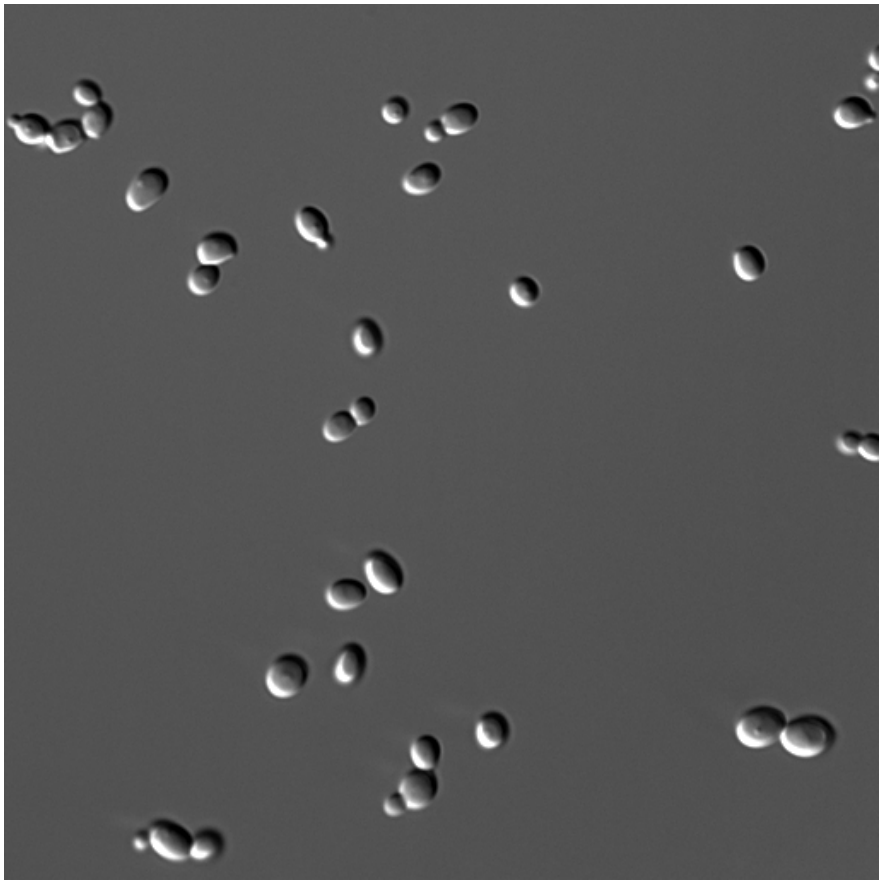


Figure 2. Differential interference contrast micrograph of budding cells of *Nakazawaea atacamensis* sp. nov. Micrograph of budding cells grown in YPD broth after 3 days at 25°C**.** Bars: 5 µm. This image was obtained using differential interference contrast (DIC) microscopy.

The novel species *N. atacamensis* exhibits distinctive phenotypic characteristics compared to other species within the genus*Nakazawaea* , as outlined in **Table S2** . *N. pomicola*ferment D-glucose, while *N. laoshanensis* is able to ferment glucose and galatactose. In contrast, *N. atacamensis* ferments both D-glucose, galactose (variable) and sucrose (Tiwari *et al.* , 2022). One notable difference is the ability of *N. atacamensis* to assimilate trehalose, while it cannot metabolize inulin and soluble starch. In contrast, *N. pomicola* and *N. laoshanensis* are capable of assimilating soluble starch. *N. laoshanensis* can assimilate DL-Lactate whereas *N. atacamensis* does not possess this capability. These phenotypic variations highlight the distinct metabolic characteristics of *N. atacamensis* when compared to other *Nakazawaea* species, emphasizing its unique physiological profile within the genus.

Nakazawaea is a genus of ascomycetous yeasts belonging to the class Pichiomycetes, order Alaninales, and family Pachysolenaceae (Groenewald *et al.* , 2023). Currently, Mycobank lists 15 valid species within the genus *Nakazawaea* . These yeasts have been isolated from diverse habitats, with a predominant association with plant materials such as apples, grapes, and fermented grape musts. They have also been found in sugar cane leaves, decayed wood, and in association with wood-feeding insects like beetles (Yamada *et al.* , 1994, Kurtzman, 2001, Kurtzman, 2011, Kurtzman & Robnett, 2014, Polburee *et al.* , 2017, Crous *et al.* , 2019, Tiwari *et al.* , 2022). The diversity of *Nakazawaea* species reflects their ability to thrive in different environments and highlights their importance in various ecological contexts. Their occurrence in plant materials and wood-associated habitats underscores their potential ecological and industrial significance,

including their potential involvement in the production of fermented foods and beverages. Recent yeast surveys in mushrooms in China and termites extended the repertoire of species in the genus *Nakazawaea* (Tiwari *et al.* , 2022) indicating a wide habitat for *Nakazawaea* species. In the present study, three strains of a novel species *N. atacamensis* were isolated from the bark and bark-exuded gum of the *N. chilensis* in the Atacama Desert, San Pedro, Chile.

3.3.2. *Nakazawaea atacamensis* whole genome sequencing

Based on the genome sequencing and assembly of the *N. atacamensis*ATA-11A-B^{T} isolate, we assessed the genome complexity and quality of the novel species. Illumina sequencing yielded approximately 7.2 x 10^5 filtered reads, providing a sequence coverage depth of 8.8X. Seven genome assemblers were compared, and the best assembly was obtained using SPAdes/Shovill (**Table S3**). The resulting genomic draft of *N. atacamensis* had a 12.4 Mbp length, with an estimated GC content of 36.7%. The *N. atacamensis* assembly consisted of 115 contigs, with 42 contigs exceeding 501 bp, accounting for 99% of the assembled sequences. The largest contig had a length of 2,070.2 Kbp, and the N50 value of the assembly was 729,094 bp, indicating the contiguity of the assembly. The details of the sequencing results and assembled contigs can be found in**Table S3** .

Next, we utilize the available genomes from the *Nakazawaea* genus to generate a phylogenetic tree. The genus *Nakazawaea* currently contains four publicly available genomes: *N. ishiwadae* , *N. holstii* , *N. peltata* and *N. ambrosiae* . The phylogenetic tree placed *N. atacamensis* together with *N. peltata*(**Figure 3**), which differs from the D1/D2 phylogeny. Differences can arise since the D1/D2 marker represents a single-locus DNA marker, thus potentially inheriting limitations associated with its use, such as missing information between sister taxa as to compared to whole-genome data (Kauserud, 2023). In general, the level of phylogenetic resolution allowed by single genes is insufficient, while the utilization of whole genome concatenated orthologous genes represents a better approximation to resolve the phylogenetic relationship between species (Brown *et al.* , 2001, Yokono *et al.* , 2018)

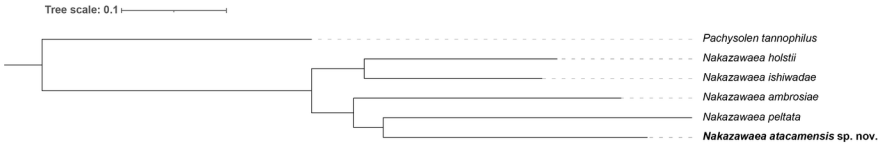


Figure 3. Maximum likelihood phylogenetic tree using whole-genome sequences. A concatenated alignment of 2,422 single-copy orthologs was used to construct the ML tree. *P. tannophilus* was used as an outgroup. Branch lengths denote amino acid substitutions per site.

To validate species discrimination, we employed the Average Nucleotide Identity (ANI) analysis across *Nakazawaea* genomes. Our examination revealed an average ANI value of 72.4% between *N. atacamensis* and the remaining genomes (**Table S4**). Consequently, this finding supports the classification of *N. atacamensis* as a novel species, consistent with the established yeast species delineation criteria (Lachance *et al.* , 2020). ANI serves as a robust parameter for demarcating species boundaries in yeasts using genome sequence data. Specifically, ANI values below 95%, which are indicative of distinct bacterial species, have been found also to be a good guideline for a group of well-defined yeast species (Lachance*et al.* , 2020).

Notably, the genome size of *N. atacamensis* is comparable to that of other *Nakazawaea* species, such as *N. ishiwadae* GDMCC 60786. Gene prediction and component analysis of the *N. atacamensis* genome using the GeneMark tool resulted in the identification of 5,394 predicted genes. Among these genes, 5,116 protein-coding genes (95%) were annotated with InterProScan (**Table S5**). To facilitate the reconstruction of the molecular network from the predicted proteins, we employed KofamKOALA and assigned KEGG Orthologs (KOs). We identified 2,782 genes involved in 385 pathways (**Table S6**). Most of the predicted genes are associated with metabolic pathways, biosynthesis of secondary metabolites, microbial metabolism in diverse environments, and biosynthesis of cofactors (**Figure 4A**). Given that *N. atacamensis* is a fermenting yeast,

we specifically focused on carbon metabolism. Our analysis revealed the presence of 68 genes encoding enzymes involved in various carbon source metabolism pathways (**Figure 4B** , **Table S6**). These pathways include glycolysis/gluconeogenesis, pyruvate metabolism, the citrate cycle (TCA), and the pentose phosphate pathway, all of which are critical for sugar fermentation through the central carbon metabolism. Additionally, other pathways, such as glycogen biosynthesis and degradation, nucleotide sugar biosynthesis, and UDP-N-acetyl-D-glucosamine biosynthesis, may also be present in *N. atacamensis* .

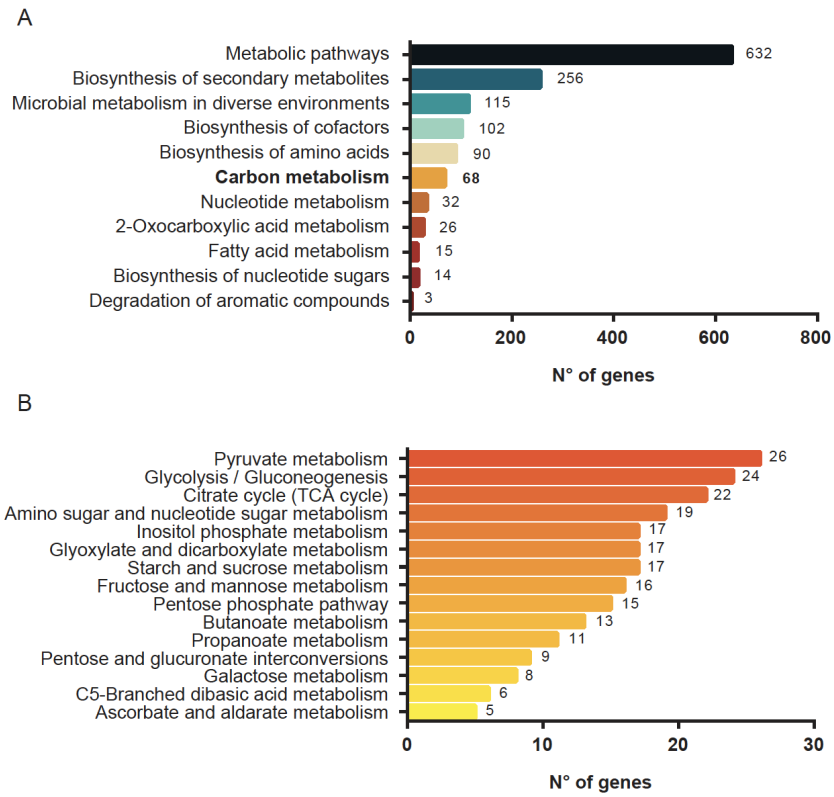


Figure 4. ***N. atacamensis*** **genome analyses. A.** KEGG distribution based on gene prediction using KofamKOALA. **B.** KEGG enrichment analysis of the carbohydrate metabolism pathways. In both figures the x-axis depicts the number of genes on each pathway identified in the *N. atacamensis* genome.

3.4 *N. atacamensis* exhibits a potential utilization in wine fermentation.

To determine the potential of *N. atacamensis* for the elaboration of alcoholic beverages, we conducted initial evaluations of biomass production under microculture conditions utilizing the ATA-11A-B^{T} strain as representative of the species. Various carbon sources and fermentation-related conditions, including glucose and fructose utilization, as well as ethanol and high glucose concentration tolerance, were examined, with the EC1118 wine *S. cerevisiae* strain used as a comparison. Overall, *N. atacamensis*ATA-11A-B^{T} showed lower growth rates (μmax) as compared to EC1118 when cultured with 2% glucose or 2% fructose (p -value < 0.05, one-way ANOVA, **Figure 5A** ,**Table S7**). When we evaluated growth under the sucrose disaccharide as a carbon source, we did not observe significant differences between the *N. atacamensis* and the *S. cerevisiae* strains, where ATA-11A-B^{T} exhibited a lower μmax than the *S. cerevisiae* wine strain (p -value > 0.05, one-way ANOVA, **Figure 5A**).

Subsequently, we subjected ATA-11A-B^{T} to ethanol and high glucose concentrations as stressors. ATA-11A-B^{T}displayed a high tolerance to ethanol, as it was able to grow in the presence of ethanol up to 8%

v/v (**Figure 5B**). However, the growth rates of ATA-11A-B^{T} under ethanol conditions were still lower compared to those of EC1118 (p -value < 0.05, one-way ANOVA). Despite this difference, the observed ethanol tolerance of ATA-11A-B^{T} highlights its potential for fermentation processes. Furthermore, the *N. atacamensis* strain exhibited a high μmax when cultured in 20% glucose, indicating its potential suitability for wine fermentation conditions. In this sense, other species from the same genus, such as *N. ishiwadae* have been previously reported to possess high ethanol tolerance and show potential for oenological applications (Ruiz *et al.* , 2019, van Wyk *et al.* , 2020).

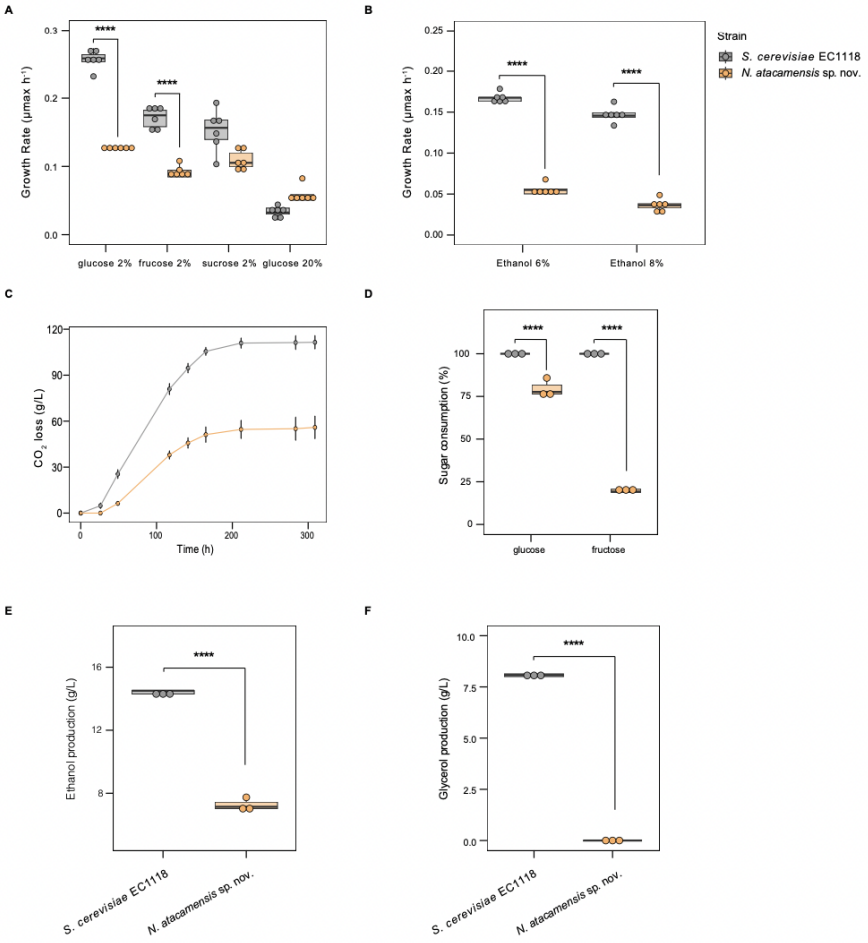


Figure 5. Fermentation phenotypes in *N. atacamensis* . **A** . Growth rate under glucose, fructose and sucrose 2% w/v, and glucose 20% w/vas carbon sources. **B** . Growth rates under medium supplemented with glucose 2% w/v and ethanol at 6 and 8% v/v. Fermentation results in Synthetic Wine Must (SWM) with a YAN concentration of 300 mg/mL.**C** . CO2 loss, **D** . Sugar consumption, **E.** Ethanol production and **F** . Glycerol production. EC1118 (grey dots) and*N. atacamensis* (orange dots).

Based on the results obtained from the microcultures, we proceeded to conduct wine fermentations using monocultures of the *N. atacamensis* type strain. To assess its fermentation capability, we measured the amount of CO_2 lost under Synthetic Wine Must (SWM) conditions, with a YAN concentration of 300 mg/mL. The fermentation assay revealed that ATA-11A-B^{T} was capable of fermenting SWM, achieving a maximum of 75,5 g/L, which was significantly lower compared to the commercially available *S. cerevisiae* control strain (p -value < 0.05, one-way ANOVA, **Figure 5C**). The *N. atacamensis* type strain exhibited incomplete fermentation (**Figure 5D**), producing ethanol levels of 7.3 ± 0.4% v/v (**Figure 5E**) and a

having a residual sugar content of 126.7 g/L, comprising 25.8 ± 7.0 g/L of glucose and 100.9 ± 2.5 g/L of fructose (**Figure 5D**). In contrast, the EC1118 control demonstrated a higher ethanol and glycerol production (**Figure 5F**), reaching up to 14.4 ± 0.2% v/v of ethanol, with only 23.6 ± 2.4 g/L of residual fructose (**Table S8**). Nevertheless, it is important to note that these values are comparable to those observed with other non-conventional yeasts used in the wine industry, such as *Torulaspora delbrueckii* ,*Metschnikowia pulcherrima* , *Starmerella bacillaris* ,*Wickerhamomyces anomalus* and *Hanseniaspora vineae*(Vejarano & Gil-Calderón, 2021). Altogether, our findings suggest the potential utilization of the novel species *N. atacamensis* in monocultures or in co-cultures with *S. cerevisiae* for wine fermentation.

DATASET

All the data generated or analyzed in this study is included in this article (supplementary information files). The strains used in this study are available on request. All fastq sequences were deposited in the National Center for Biotechnology Information (NCBI) as a Sequence Read Archive under the BioProject accession number PRJNA991564 (www.ncbi.nlm.nih.gov/bioproject/ PRJNA991564).

ACKNOWLEDGEMENTS

FC acknowledges ANID FONDECYT [1220026] and ANID - Programa Iniciativa Científica Milenio NCN2021_050. FC, RGA, TM and FPD are grateful to ANID - Programa Iniciativa Científica Milenio - ICN17_022, FPD also acknowledges AFOREST Nucleus NCS2022_24. RG acknowledges ANID – Programa Iniciativa Científica Milenio, Center for Genome Regulation - ICN2021_044. MA acknowledges Proyecto DICYT [022343CR], Vicerrectoría de Investigación, Desarrollo e Innovación, USACH. CAR was funded by "INCT Yeasts: Biodiversity, preservation and biotechnological innovation", funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil, grant #406564/2022-1, and CNPq process numbers 0457499/2014-1, 313088/2020-9, and 408733/2021; and Fundação do Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG, process numbers APQ-01525-14, APQ-01477-13 and APQ-02552-15). TM is supported by ANID FONDECYT POSTDOCTORADO [grant 3220801]. We acknowledge Marc-André Lachance for providing additional CBS 5808 phenotypic data.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Aponte-Santamaría C, Fischer G, Båth P, Neutze R & de Groot BL (2017) Temperature dependence of protein-water interactions in a gated yeast aquaporin. *Scientific Reports* **7** : 1-14.

Astudillo L, Schmeda-Hirschmann G, Herrera JP & Cortes M (2000) Proximate composition and biological activity of Chilean Prosopis species. *Journal of the Science of Food and Agriculture***80** : 567-573.

Azua-Bustos A, Urrejola C & Vicuna R (2012) Life at the dry edge: microorganisms of the Atacama Desert. *FEBS Lett* **586** : 2939-2945.

Azua-Bustos A, Fairen AG, Gonzalez-Silva C*, et al.* (2018) Unprecedented rains decimate surface microbial communities in the hyperarid core of the Atacama Desert. *Scientific Reports***8** : 16706.

Brown JR, Douady CJ, Italia MJ, Marshall WE & Stanhope MJ (2001) Universal trees based on large combined protein sequence data sets.*Nat Genet* **28** : 281-285.

Buzzini P, Turchetti B & Yurkov A (2018) Extremophilic yeasts: the toughest yeasts around? *Yeast* **35** : 487-497.

Canonico L, Solomon M, Comitini F, Ciani M & Varela C (2019) Volatile profile of reduced alcohol wines fermented with selected non-Saccharomyces yeasts under different aeration conditions. *Food Microbiology* **84** : 103247.

Carrasco-Puga G, Diaz FP, Soto DC, Hernandez-Castro C, Contreras-Lopez O, Maldonado A, Latorre C & Gutierrez RA (2021) Revealing hidden plant diversity in arid environments. *Ecography* **44** : 98-111.

Crous PW, Carnegie AJ, Wingfield MJ, Sharma R, Mughini G, Noordeloos ME, Santini A, Shouche YS, Bezerra JD & Dima B (2019) Fungal Planet description sheets: 868–950. *Persoonia: Molecular Phylogeny and Evolution of Fungi* **42** : 291.

Cubillos FA, Gibson B, Grijalva-Vallejos N, Krogerus K & Nikulin J (2019) Bioprospecting for brewers: Exploiting natural diversity for naturally diverse beers. *Yeast* **36** : 383-398.

Drees KP, Neilson JW, Betancourt JL, Quade J, Henderson DA, Pryor BM & Maier RM (2006) Bacterial community structure in the hyperarid core of the Atacama Desert, Chile. *Appl Environ Microbiol* **72** : 7902-7908.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32** : 1792-1797.

Eizaguirre JI, Peris D, Rodriguez ME, Lopes CA, De Los Rios P, Hittinger CT & Libkind D (2018) Phylogeography of the wild Lager-brewing ancestor (Saccharomyces eubayanus) in Patagonia. *Environmental microbiology* **20** : 3732-3743.

Eshel G, Araus V, Undurraga S*, et al.* (2021) Plant ecological genomics at the limits of life in the Atacama Desert. *Proceedings of the National Academy of Sciences* **118** : e2101177118.

Ewing SA, Sutter B, Owen J, Nishiizumi K, Sharp W, Cliff SS, Perry K, Dietrich W, McKay CP & Amundson R (2006) A threshold in soil formation at Earth's arid–hyperarid transition. *Geochimica et Cosmochimica Acta* **70** : 5293-5322.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *evolution* **39** : 783-791.

Finstad KM, Probst AJ, Thomas BC, Andersen GL, Demergasso C, Echeverria A, Amundson RG & Banfield JF (2017) Microbial Community Structure and the Persistence of Cyanobacterial Populations in Salt Crusts of the Hyperarid Atacama Desert from Genome-Resolved Metagenomics. *Front Microbiol* **8** : 1435.

Groenewald M, Hittinger CT, Bensch K*, et al.* (2023) A genome-informed higher rank classification of the biotechnologically important fungal subphylum Saccharomycotina. *Studies in Mycology* .

Hall T, Biosciences I & Carlsbad C (2011) BioEdit: an important software for molecular biology. *GERF Bull Biosci* **2** : 60-61.

Hillis DM & Bull JJ (1993) An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis.*Systematic biology* **42** : 182-192.

Houston J & Hartley AJ (2003) The central Andean west-slope rainshadow and its potential contribution to the origin of hyper-aridity in the Atacama Desert. *International Journal of Climatology* **23** : 1453-1464.

Kauserud H (2023) ITS alchemy: On the use of ITS as a DNA marker in fungal ecology. *Fungal Ecology* **65** : 101274.

Kumar S, Stecher G, Li M, Knyaz C & Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* **35** : 1547-1549.

Kurtzman CP (2001) Six new anamorphic ascomycetous yeasts near Candida tanzawaensis. *FEMS Yeast Research* **1** : 177-185.

Kurtzman CP (2011) Nakazawaea Y. Yamada, Maeda & Mikata (1994). *The Yeasts,* p.ˆpp. 637-639. Elsevier.

Kurtzman CP & Robnett CJ (1998) Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. *Antonie Van Leeuwenhoek***73** : 331-371.

Kurtzman CP & Robnett CJ (2014) Description of Kuraishia piskuri fa, sp. nov., a new methanol assimilating yeast and transfer of phylogenetically related Candida species to the genera Kuraishia and Nakazawaea as new combinations. *FEMS yeast research* **14** : 1028-1036.

Lachance M-A (2012) In defense of yeast sexual life cycles: The forma asexualis-an informal proposal. *Yeast Newsletter* **61** : 24.

Lachance M-A, Lee DK & Hsiang T (2020) Delineating yeast species with genome average nucleotide identity: a calibration of ANI with haplontic, heterothallic Metschnikowia species. *Antonie van Leeuwenhoek***113** : 2097-2106.

Lachance MA (2018) C. P. Kurtzman's evolving concepts of species, genus and higher categories. *FEMS Yeast Res* **18** .

Lachance MA, Bowles JM, Starmer WT & Barker JS (1999) Kodamaea kakaduensis and Candida tolerans, two new ascomycetous yeast species from Australian Hibiscus flowers. *Can J Microbiol* **45** : 172-177.

Lee I, Ouk Kim Y, Park SC & Chun J (2016) OrthoANI: An improved algorithm and software for calculating average nucleotide identity.*Int J Syst Evol Microbiol* **66** : 1100-1103.

Libkind D, Hittinger CT, Valerio E, Goncalves C, Dover J, Johnston M, Goncalves P & Sampaio JP (2011) Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast.*Proc Natl Acad Sci U S A* **108** : 14539-14544.

McGinnis S & Madden TL (2004) BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res***32** : W20-25.

McKay CP, Friedmann EI, Gomez-Silva B, Caceres-Villanueva L, Andersen DT & Landheim R (2003) Temperature and moisture conditions for life in the extreme arid region of the Atacama desert: four years of observations including the El Nino of 1997-1998. *Astrobiology* **3** : 393-406.

Moncalvo J-M, Lutzoni FM, Rehner SA, Johnson J & Vilgalys R (2000) Phylogenetic Relationships of Agaric Fungi Based on Nuclear Large Subunit Ribosomal DNA Sequences. *Systematic Biology* **49** : 278-305.

Morais PB, Pagnocca FC & Rosa CA (2006) Yeast Communities in Tropical Rain Forests in Brazil and other South American Ecosystems.*Biodiversity and Ecophysiology of Yeasts,* (Peter G & Rosa C, eds.), p.ˆpp. 461-484. Springer Berlin Heidelberg, Berlin, Heidelberg.

Mozzachiodi S, Bai F-Y, Baldrian P*, et al.* (2022) Yeasts from temperate forests. *Yeast* **39** : 4-24.

Navarro-Gonzalez R, Rainey FA, Molina P, Bagaley DR, Hollen BJ, de la Rosa J, Small AM, Quinn RC, Grunthaner FJ & Caceres L (2003) Mars-like soils in the Atacama Desert, Chile, and the dry limit of microbial life.*Science* **302** : 1018-1021.

Nespolo RF, Villarroel CA, Oporto CI, Tapia SM, Vega-Macaya F, Urbina K, De Chiara M, Mozzachiodi S, Mikhalev E & Thompson D (2020) An Out-of-Patagonia migration explains the worldwide diversity and distribution of Saccharomyces eubayanus lineages. *PLoS Genetics***16** : e1008777.

Pardo O. PJ (2015) *Chile. Bebidas Fermentadas Prehispanicas* . Ediciones Parina EIRL.

Peris D, Langdon QK, Moriarty RV, Sylvester K, Bontrager M, Charron G, Leducq J-B, Landry CR, Libkind D & Hittinger CT (2016) Complex ancestries of lager-brewing hybrids were shaped by standing variation in the wild yeast Saccharomyces eubayanus. *PLoS genetics***12** : e1006155.

Polburee P, Lertwattanasakul N, Limtong P, Groenewald M & Limtong S (2017) Nakazawaea todaengensis fa, sp. nov., a yeast isolated from a peat swamp forest in Thailand. *International Journal of Systematic and Evolutionary Microbiology* **67** : 2377-2382.

Pulschen AA, Rodrigues F, Duarte RT, Araujo GG, Santiago IF, Paulino-Lima IG, Rosa CA, Kato MJ, Pellizari VH & Galante D (2015) UV-resistant yeasts isolated from a high-altitude volcanic area on the Atacama Desert as eukaryotic models for astrobiology.*Microbiologyopen* **4** : 574-588.

Ruiz J, Ortega N, Martin-Santamaria M, Acedo A, Marquina D, Pascual O, Rozes N, Zamora F, Santos A & Belda I (2019) Occurrence and enological properties of two new non-conventional yeasts (Nakazawaea ishiwadae and Lodderomyces elongisporus) in wine fermentations. *International Journal of Food Microbiology* **305** : 108255.

Schulze-Makuch D, Wagner D, Kounaves SP*, et al.* (2018) Transitory microbial habitat in the hyperarid Atacama Desert. *Proc Natl Acad Sci U S A* **115** : 2670-2675.

Sciammaro LP, Ribotta DP & Puppo MC (2016) Traditional Food Products from Prosopis sp. Flour. *Traditional Foods: General and Consumer Aspects,* (Kristbergsson K & Oliveira J, eds.), p.ˆpp. 209-216. Springer US, Boston, MA.

Sciammaro LP, Puppo MC & Voget C (2022) Chapter 25 - "Aloja": A pre-Hispanic fermented beverage from Prosopis alba pods. *Prosopis as a Heat Tolerant Nitrogen Fixing Desert Food Legume,* (Puppo MC & Felker P, eds.), p.ˆpp. 341-351. Academic Press.

STACKEBRANDT E & GOEBEL BM (1994) Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology. *International Journal of Systematic and Evolutionary Microbiology* **44** : 846-849.

Tiwari S, Behera BC & Baghela A (2022) Nakazawaea odontotermitis f.a., sp. nov., a novel yeast isolated from the gut of Odontotermes horni in India. *Arch Microbiol* **204** : 237.

Tiwari S, Behera BC & Baghela A (2022) N akazawaea odontotermitis fa, sp. nov., a novel yeast isolated from the gut of Odontotermes horni in India. *Archives of Microbiology* **204** : 237.

van Wyk N, Pretorius IS & von Wallbrunn C (2020) Assessing the Oenological Potential of Nakazawaea ishiwadae, Candida railenensis and Debaryomyces hansenii Strains in Mixed-Culture Grape Must Fermentation with Saccharomyces cerevisiae. *Fermentation* **6** : 49.

Vejarano R & Gil-Calderon A (2021) Commercially Available Non-Saccharomyces Yeasts for Winemaking: Current Market, Advantages over Saccharomyces, Biocompatibility, and Safety. *Fermentation***7** : 171.

Vilgalys R & Hester M (1990) Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA from several Cryptococcus species. *J Bacteriol* **172** : 4238-4246.

Villarreal P, Quintrel PA, Olivares-Munoz S, Ruiz JJ, Nespolo RF & Cubillos FA (2022) Identification of new ethanol-tolerant yeast strains with fermentation potential from central Patagonia. *Yeast***39** : 128-140.

Vu D, Groenewald M, Szoke S*, et al.* (2016) DNA barcoding analysis of more than 9 000 yeast isolates contributes to quantitative thresholds for yeast species and genera delimitation. *Stud Mycol***85** : 91-105.

Wei XY, Zhu HY, Song L, Zhang RP, Li AH, Niu QH, Liu XZ & Bai FY (2022) Yeast Diversity in the Qaidam Basin Desert in China with the Description of Five New Yeast Species. *J Fungi (Basel)* **8** .

White TJ, Bruns T, Lee S & Taylor J (1990) Amplification and Direct Sequencing of Fungal Ribosomal Rna Genes for Phylogenetics. *PCR Protocols,* (Innis MA, Gelfand DH, Sninsky JJ & White TJ, eds.), p.ˆpp. 315-322. Academic Press, San Diego.

Wierzchos J, Camara B, de Los Rios A, Davila AF, Sanchez Almazo IM, Artieda O, Wierzchos K, Gomez-Silva B, McKay C & Ascaso C (2011) Microbial colonization of Ca-sulfate crusts in the hyperarid core of the Atacama Desert: implications for the search for life on Mars.*Geobiology* **9** : 44-60.

Yamada Y, Maeda K & Mikata K (1994) The phylogenetic relationships of the hat-shaped ascospore-forming, nitrate-assimilating Pichia species, formerly classified in the genus Hansenula Sydow et Sydow, based on the

partial sequences of 18S and 26S ribosomal RNAs (Saccharomycetaceae): the proposals of three new genera, Ogataea, Kuraishia, and Nakazawaea.*Bioscience, biotechnology, and biochemistry* **58** : 1245-1257.

Yokono M, Satoh S & Tanaka A (2018) Comparative analyses of whole-genome protein sequences from multiple organisms. *Scientific Reports* **8** : 6800.

FIGURE LEGENDS

Figure 1. Neighbour-joining phylogram showing the placement of *Nakazawaea atacamensis* sp. nov sp. nov. within the genus . The tree was built based on the sequences of the D1/D2 domains of the LSU rRNA gene. The alignment was performed with MUSCLE. *Pachysolen tannophilus* was used as an outgroup. The distance metric is the number of substitutions. Bar, Substitutions per site.

Figure 2. Differential interference contrast micrograph of budding cells of *Nakazawaea atacamensis* sp. nov. Micrograph of budding cells grown in YPD broth after 3 days at 25degC**.** Bars: 5 μm. This image was obtained using differential interference contrast (DIC) microscopy.

Figure 3. Maximum likelihood phylogenetic tree using whole-genome sequences. A concatenated alignment of 2,422 single-copy orthologs was used to construct the ML tree. *P. tannophilus* was used as an outgroup. Branch lengths denote amino acid substitutions per site.

Figure 4. *N. atacamensis* genome analyses. A. KEGG distribution based on gene prediction using KofamKOALA. **B.** KEGG enrichment analysis of the carbohydrate metabolism pathways. In both figures the x-axis depicts the number of genes on each pathway identified in the *N. atacamensis* genome.

Figure 5. Fermentation phenotypes in *N. atacamensis* . **A** . Growth rate under glucose, fructose and sucrose 2% w/v, and glucose 20% w/vas carbon sources. **B** . Growth rates under medium supplemented with glucose 2% w/v and ethanol at 6 and 8% v/v. Fermentation results in Synthetic Wine Must (SWM) with a YAN concentration of 300 mg/mL. **C** . CO2 loss, **D** . Sugar consumption, **E.** Ethanol production and **F** . Glycerol production. EC1118 (grey dots) and *N. atacamensis*(orange dots).

SUPPLEMENTARY INFORMATION

TABLE LEGENDS

Table S1. Number of yeast isolates obtained from tree samples from the Atacama Desert.

Table S2 . Phenotypic characterisation of *N. atacamensis.*

Table S3 . De-Novo Genome Assemblers statistics for *N. atacamensis.*

Table S4. OrthoANI results across *Nakazawaea* species.

Table S5 . Gene predictions obtained from whole-genome sequencing.

Table S6 . KEGG pathways in *N. atacamensis* .

Table S7 . Growth rates under fermentation-related conditions.

Table S8 . Total CO_2 loss, residual sugars, ethanol and glycerol production in the Synthetic Wine Must (SWM) fermentation.

FIGURE LEGENDS

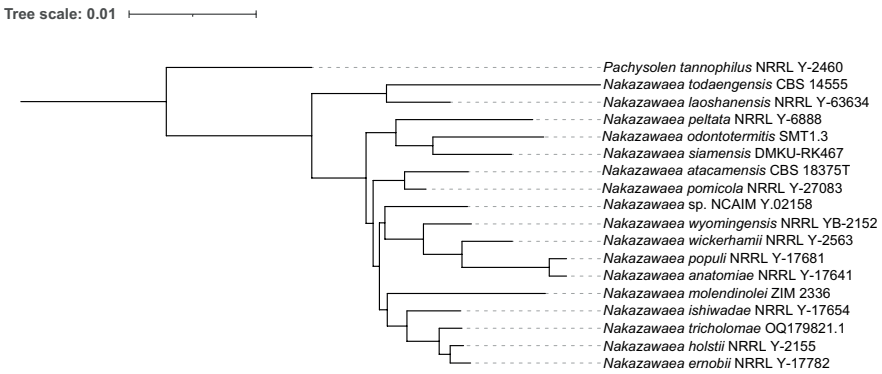
Figure S1 . **ITS region sequence alignment of *N. atacamensis* sp. nov., *N. peltata*, *N. siamensis* and *N. odontotermitis*** . The strain details are *N. peltata* CBS 5576, *N. siamensis* CBS 12569 and *N. odontotermitis* SMT1.3.

Figure S2. D1/D2 domains of the LSU rRNA gene sequence alignment of three *N. atacamensis* strains. The GenBank accession numbers for the D1/D2 domains of the LSU rRNA gene are

N. atacamensis ATA-11A-B^{T} (OP293331), *N. atacamensis* ATA-11A-B^{T} (OP293332) and *N. atacamensis* ATA-13E-S (OP293333).

Figure S3. SSU rRNA gene sequence alignment of three *N. atacamensis* strains. The GenBank accession numbers for SSU rRNA gene are *N. atacamensis*ATA-11A-B^{T} (OP293328), *N. atacamensis*ATA-11A-B^{T} (OP293329) and *N. atacamensis*ATA-13E-S (OP293330).

Figure S4. Pairwise identity matrix analysis. The pairwise comparison was made considering the SSU rRNA gene, the ITS region, and the D1/D2 domains of the LSU rRNA gene from various *Nakazawaea*species. The purple colours in the identity matrix indicate high sequence identity, while light blue colours low sequence identity.



figures/231004-Figure-2/231004-Figure-2-eps-converted-to.pdf