Chimeric Sex-Determining Chromosomal Regions and Dysregulation of Cell-Type Identity in a Sterile Zygosaccharomyces Allodiploid Yeast

Melissa Bizzarri, Paolo Giudici, Stefano Cassanelli, Lisa Solieri *

Department of Life Sciences, University of Modena and Reggio Emilia, Via Amendola 2, 42122, Reggio Emilia, Italy

* lisa.solieri@unimore.it

Abstract

Allodiploidization is a fundamental yet evolutionarily poorly characterized event, which impacts genome evolution and heredity, controlling organismal development and polyploid cell-types. In this study, we investigated the sex determination system in the allodiploid and sterile ATCC 42981 yeast, a member of the Zygosaccharomyces rouxii species complex, and used it to study how a chimeric mating-type gene repertoire contributes to hybrid reproductive isolation. We found that ATCC 42981 has 7 MAT-like (MTL) loci, 3 of which encode α-idiomorph and 4 encode a-idiomorph. Two phylogenetically divergent MAT expression loci were identified on different chromosomes, accounting for a hybrid a/α genotype. Furthermore, extra a-idiomorph-encoding loci (termed MTLa copies 1 to 3) were recognized, which shared the same MATα1 ORFs but diverged for MATα2 genes. Each MAT expression locus was linked to a HML silent cassette, while the corresponding HMR loci were located on another chromosome. Two putative parental sex chromosome pairs contributed to this unusual genomic architecture: one came from an as-yet-undescribed taxon, which has the NCYC 3042 strain as a unique representative, while the other did not match any MAT-HML and HMR organizations previously described in Z. rouxii species. This chimeric rearrangement produces two copies of the HO gene, which encode for putatively functional endonucleases essential for mating-type switching. Although both a and α coding sequences, which are required to obtain a functional cell-type a1-α2 regulator, were present in the allodiploid ATCC 42981 genome, the transcriptional circuit, which regulates entry into meiosis in response to meiosis-inducing salt stress, appeared to be turned off. Furthermore, haploid and α-specific genes, such as MATα1 and HO, were observed to be actively transcribed and up-regulated under hypersaline stress. Overall, these evidences demonstrate that ATCC 42981 is unable to repress haploid α-specific genes and to activate meiosis in response to stress. We argue that sequence divergence within the chimeric a1-α2 heterodimer could be involved in the generation of negative epistasis, contributing to the allodiploid sterility and the dysregulation of cell identity.
Introduction

Ploidy variation has played a major role in the evolution of many extant eukaryotic lineages [1]. In yeasts, numerous studies have demonstrated how variations in the ploidy state took place frequently during evolutionary history [2] [3]. Allopolyploid offspring often have strong selective disadvantages due to their sterility [4]. However, in some instances, the increased genome size and complexity of allopolyploids may enhance heterosty and/or adaptive flexibility [5], particularly at the edges of the ancestral species’ range, where they are more likely to encounter stress [6] [7]; [8]. Life history models note that there is an intricate interplay between ploidy variation and alterations of mating, meiosis and sporulation patterns [9]. Consequently, the ploidy state affects the genetic composition at sex-determining loci, giving rise to the ploidy-dependent initiation of dedicated transcriptional programs [10].

In the well-studied model organism Saccharomyces cerevisiae, the meiosis of diploid cells is triggered by environmental cues, such as nutrient depletion, and gives rise to four haploid meiotic spores with two distinct mating-types (MATa and MATα). Haploid spores eventually re-establish diploid MATa/MATα lines by one of three processes: (1) by mating with their own mitotic daughter cells after switching their mating-type in a process catalysed by the endonuclease HO (HOMothallism); (2) by mating with another spore created by the same meiotic event (intratetrad mating); or, more rarely, (3) by mating with an unrelated individual (out-crossing) [11]. The α1-α2 protein heterodimer is one of the master regulators of cell identity and sexual development (as reviewed in [12] and [13], respectively). The α1 and α2 proteins are homeodomain transcriptional factors encoded by the MATα1 and MATα2 genes at the active MATα and MATα loci, respectively. In MATα/MATα diploid cells, α2 interacts with α1 to bind DNA as a heterodimer and transcriptionally repress mating genes, preventing polyploidy and aneuploidy. Furthermore, the α1-α2 heterodimer inhibits the expression of haploid-specific genes (h-sgs) and activates the developmental switch from mitosis to meiosis under appropriate stress stimuli [14] [15]. Two h-sg targets of negative regulation by the α1-α2 heterodimer are MATα1 (encoding the transcription factor that positively regulates the α-specific genes) and HO genes (encoding an endonuclease, which cleaves a specific DNA sequence at the MAT locus during the first step of mating-type switching) [16]. Furthermore, α1-α2 heterodimer controls the sexual development by inhibiting the RME1 (Repressor of Meiosis 1) gene. In haploid cells, the RME1 transcriptional factor prevents the expression of the IME1 (Inducer of Meiosis 1) gene [17] [18] and positively regulates the transcription of adhesion-specific genes, such as the FLO11 gene, in response to nutrient depletion [19]. In diploid cells, the α1-α2 heterodimer binds to the RME1 gene promoter, repressing its transcription and thereby relieving IME1 gene repression. The α1-α2 repressor complex directly controls entry into meiosis through the regulation of the IME4 gene, which is required for the full expression of IME1 [20]. The α1-α2 heterodimer negatively regulates the antisense (AS) long non-coding RNA (lncRNA) [21] [22]. This lncRNA (also termed as Regulator of Meiosis 2 or RME2) is transcribed in haploid cells and blocks the expression of the IME4 gene. In diploid cells, the α1-α2 complex represses RME2 expression, allowing sense (S)-IME4 to be transcribed under starvation conditions.

Yeasts of the non-whole genome duplication (non-WGD) Zygosaccharomyces rouxii species complex are adapted to grow in food with high solute concentrations and frequently experience variations in ploidy, resulting in different modes of reproduction [23]. This complex includes the haploid heterothallic or homothallic (self-fertile) Z. rouxii species, which undergoes mating and subsequent meiosis under salt stimuli [23] [24]; the diploid Zygosaccharomyces sapae species, which reproduces mainly by clonality and ascospores are rarely observed [25]; and a group of aneuploid/alldiploid strains of unclear taxonomical position (termed mosaic lineage) [23].
Within the latter, the allodiploid ATCC 42981 strain has been extensively studied for its ability to withstand high concentrations of alkali metal cations and for its capability to produce glycerol under salt stress [27][28]. From a molecular point of view, the ATCC 42981 genome displays rDNA heterogeneity [26][29], additional chromosomes compared to Z. rouxii [28], and diploid DNA [26]. Gordon and Wolfe [29] showed that the ATCC 42981 genome contains two partially divergent complements, namely the T- and P- subgenomes, which could arise from a recent allodiploidization event between Z. rouxii carrying the T- subgenome and another not-yet-recognized species harbouring the P- subgenome (so far represented by the unique strain NCYC 3042 and referred to as Zygosaccharomyces pseudorouxii nom. inval. by James et al. [30]). Despite having a DNA diploid content, the ATCC 42981 strain has not been observed to undergo meiosis under different standard and stress growth conditions [25].

Because gene information retained at the MAT expression loci is essential to ensure appropriate haploid/diploid cell-type identity and functional cell development, several efforts have recently been made to characterize MAT loci in haploid Z. rouxii [31] and diploid Z. sapae [32], but not in the allodiploid ATCC 42981 strain. Like S. cerevisiae, haploid Z. rouxii wild strains possess a three-cassette system consisting of MATa or MATα expression loci and two silent cassettes of both idiomorphs, HMR and HML, respectively, which act as donor sequences during the mating-type switching presumably catalysed by HO endonuclease [33][34]. S. cerevisiae maintains both HML and HMR cassettes at different locations on the same chromosome harbouring the MAT expression locus, whereas Z. rouxii only conserves the MAT/HML linkage on chromosome C and the HMR locus on chromosome F [34]. This structural organization implies that, differently from S. cerevisiae, Z. rouxii exploits the ectopic recombination between non-homologous chromosomes to switch the mating-type. Congruently, Watanabe et al. [31] found that sex chromosomes represent hyper-mutational hotspots, with flanking regions of MAT, HML and HMR cassettes remaining highly variable among Z. rouxii haploid strains.

Compared to Z. rouxii, the Z. sapae diploid species has one more copy of the HO gene and displays an unusual αααα genotype, with a redundant number of divergent MATα loci but without any HMR silent locus. This complex architecture of sex-determining chromosomal regions prevents mating-type switching due to the lack of the HMR cassette and hampers sexual development due to the imbalance in divergent mating-type genes.

Here, we determined the structure and functions of the three-locus sex-determining system in the allodiploid ATCC 42981 strain. The transcriptional profiling of these genes under meiotic-inducing salt stress revealed how the hybrid genetic configuration at the MAT loci contributes to allodiploid sterility. To the best of our knowledge, our report presents the first evidence that a chimeric sex-determination system accounts for the incomplete silencing of the h-sg program and contributes to the prevention of switching from mitosis to meiosis in allodiploid yeasts.

Materials and Methods

Strains, culture conditions and mating test

The Zygosaccharomyces strains used in this work are listed in Table 1. Strains were routinely cultured at 28°C in YPD (1% w/v yeast extract, 1% w/v peptone, 2% w/v dextrose) medium with or without 15% (w/v) agar and stored at 4°C. For long-term preservation, strains were stored at -80°C in YPD medium containing 25% glycerol (v/v) as a cryopreservative. To increase the probability of observing conjugated or un-conjugated spore-containing ascii, sporulation was tested by inoculating the early stationary phase culture of ATCC 42981 on five different media [YPD, YPDAl, malt extract agar (MEA; Difco), MEA supplemented with 6% (w/v) NaCl (6%NaCl-MEA), and YNB5%GNaCl (1% w/v yeast extract, 5% w/v dextrose, 6.7 g/l yeast nitrogen base, 2.0 M NaCl)] for 3 weeks.
To study sexual compatibility, 2-to-4-day-old cultures of the ATCC 42981 strain were incubated alone or in a mixture with mating-tester strains \( Z.\ rouxii \) CBS 4837 (mating-type \( a \)) or CBS 4838 (mating-type \( \alpha \)) in both MEA and 6%NaCl-MEA media at 27°C for 3 weeks. Samples were examined microscopically every week using phase-contrast optics to detect conjugation.

The YNB5%G [1% w/v yeast extract, 5% w/v dextrose, 6.7 g/l yeast nitrogen base (Difco)] and the YNB5%GNaCl (1% w/v yeast extract, 5% w/v dextrose, 6.7 g/l yeast nitrogen base, 2.0 M NaCl) media were used for RNA extraction from cells grown under unstressed and salt-stressed conditions, respectively.

**PCR conditions and sequencing**

Genomic DNA (gDNA) extraction was performed by a phenol-based method from stationary grown cells after mechanical lysis according to Hoffman and Winston [35]. The quantity of DNA was evaluated spectrophotometrically using a NanoDrop ND-1000 device (Thermo Scientific). All PCR reactions were performed on a T100 Thermalcycler (Bio-Rad) in a 25-μl reaction volume containing 200 ng of template gDNA. For PCR amplification < 2 kb \( \text{Taq DNA polymerase} \) (Takara, Japan) and for PCR amplification \( \geq 2 \) kb \( \text{Phusion Hot Start II Polymerase} \) (Thermo Scientific, Waltham, MA) along with buffer HF (5x) or LA Taq DNA polymerase (Takara, Japan) with GC buffer I (2x), were used according to the manufacturer’s instructions. All primers used in this study are listed in S1 Table; primers were designed with Primer 3 software (http://primer3.sourceforge.net/) and provided by either MWG (Heidelberg, Germany) or BMR Genomics (Padova, Italy). The PCR products were resolved on 1.2% (w/v) agarose gels stained with ethidium bromide and their size was estimated by comparison with 100 bp or 1 kb DNA Ladder Plus (Fermentas) as molecular size markers. PCR products were purified using the DNA Clean & Concentrator™-5 (DCC™-5) Kit (Zymo Research) according to the manufacturer’s instructions. Moreover, when required, DNA fragments were purified from 1% agarose gels using the Gene JET Gel Extraction Kit (Thermo Scientific, Waltham, MA). Finally, all PCR products were sequenced on both strands through a DNA Sanger sequencing process performed by either MWG (Heidelberg, Germany) or BMR Genomics (Padova, Italy).

**Genomic walking procedure**

The overall strategy for determining mating-type-like (\( MTL \)) loci is depicted in S1 Fig (panel A). Briefly, three \( MATa \) copy-specific primer pairs were designed on the \( ZsMTLa \) copies 1, 2 and 3 cassettes previously cloned in \( Z.\ sapae \) and used to clone the corresponding \( MATa1 \) (partial) and \( MATa2 \) (full-length) coding regions in the ATCC 42981 genome (S1 Fig and S1 Table).

### Table 1. Details of strains used in the present study.

Ploidy data were obtained from Solieri et al. [23] [26], while the genotype of strain ABT301\(^T\) at the active \( MAT \) loci was obtained from Solieri et al. [32].

| Strains       | Other collections | Source          | Current taxonomical positions | Mating-type/thallism            | Spore       | Ploidy ratio |
|---------------|-------------------|-----------------|-------------------------------|---------------------------------|-------------|--------------|
| CBS 732\(^T\) | NCYC 568, NRRL Y-229 | Grape must      | \( Z.\ rouxii \)               | \( MATa/homothallic \)          | -           | 1.3          |
| CBS 4837      | NYC 1682, NRRL Y-2547 | Miso            | mosaic lineage                 | \( MATa/heterothallic \)         | +           | 1.96         |
| CBS 4838      | NRRL Y-2584       | Miso            | mosaic lineage                 | \( MATa/heterothallic \)         | +           | 1.90         |
| ATCC42981     | -                 | Miso            | mosaic lineage                 | nd                              | nd          | 2.1          |
| ABT301\(^T\)  | CBS 12607, MUCL54092 | TBV             | \( Z.\ sapae \)               | \( aaoMATa/MATa/heterothallic \) | +           | 2.0          |
| NCYC 3042     | CBS 9951          | Soft drink      | \( Z.\ pseudorouxii \) nom. inval. | nd                              | nd          | nd           |

nd, not determined; TBV, Traditional Balsamic Vinegar.

doi:10.1371/journal.pone.0152558.t001
Table). Complete MATα1 loci were obtained in a second round of PCR walking using reverse primers (A, B, C and A_D) built on the 3’-end regions potentially flanking the MAT and HML/HMR cassettes [31] [32] (S1 Fig). Similarly, three primer pairs were used to amplify putative MTLα loci in the ATCC 42981 strain (S1 Fig, panel B). Two primer pairs were designed to specifically amplify the ZsMATα1 and ZsMATα2 genes, respectively, whereas a third primer pair was built to completely include MATα1 ORF and partially MATα2 ORFs, respectively. In order to further extend the MATα2 coding DNA sequence, a second round of PCR walking was performed by combining a MATα2-specific reverse primer and all the available forward primers (1, 2 and 3) built on the 5’-end regions flanking the MAT and HML/HMR cassettes [31] (S1 Fig). The DNA regions flanking MTLα and MTLα loci were characterized through the seminested and direct PCR approaches as previously reported [32] (S2 Fig).

The overview of the strategy employed for ATCC 42981 HO gene characterization is summarized in S3 Fig, and a detailed primer list is given in S1 Table.

Bioinformatics analyses

Sequences were assembled and edited using DNAStar Software (DNASTAR, Inc. Madison, Wisconsin USA). Multiple nucleotide and amino acid sequence alignments were performed using Clustal W2 [36]. Searches for nucleotide and protein sequence homologs were carried out in the GenBank database with Blastn and Blastp algorithms, respectively [37]. Phylogenetic analysis was conducted on aa sequences using MEGA6 [38]. The phylogenetic relationship was inferred using the neighbour-joining (NJ) method. Support percentages for the nodes of NJ-trees were computed using bootstrapping analysis with 1,000 replications and were shown next to the branches when ≥ 60%. For domain identification, Pfam-searches [39] were run on http://pfam.sanger.ac.uk/. Structure predictions were obtained with Jpred3 [40] and validated according to Martin et al. [41]. Sequences were submitted to the EMBL/GenBank databases under the accession numbers from KT598024 to KT598027 and from KT694298 to KT6942302.

PFGE-Southern blotting assays

Chromosomal DNA preparation in plug, pulsed-field gel electrophoresis (PFGE), and Southern blotting assays were performed as previously reported [26] [32]. Primers engaged in probe synthesis for Southern blot analysis are listed in S1 Table.

Subgenome assignment of mating-type and HO gene copies

PCR subgenome genotyping was performed using gDNA extracted from Z. pseudorouxi NCYC 3042 as a T- subgenome template. PCR amplification of all MATα1, MATα2, MATα1, MATα2, and HO gene copies present in ATCC 42981 genome was carried out with the primer pairs listed in S1 Table. To avoid false negative results, we tested two alternative primer pairs for each target gene. ATCC 42981 gDNA was also included as a positive control. Z. rouxi CBS 7321 gDNA was not included in PCR reactions as one of the two putative parental genetic complements because its genome project was available.

RNA extraction and RT-PCR

Zygosaccharomyces cells were pre-cultured in YPD medium for 24 h at 28°C under shaking conditions (150 rpm), washed in physiological solution (9 g/l NaCl), and used to inoculate two sets of baffled Erlenmeyer flasks (E-flasks) containing 70 ml of YNB5%G (standard growth condition) and YNB5%GNaCl (hyperosmotic growth condition) media, respectively (initial OD600nm 0.02–0.04). Inoculated E-flasks were incubated at 28°C under shaking conditions
(150 rpm) and yeast growth was spectrophotometrically monitored at 600 nm two times a day. Cells at the stationary phase (no change in OD measurement detected in at least three consecutive readings) were harvested and frozen at -80°C. RNA extractions were carried out using the ZR Fungal/Bacterial RNA MicroPrep™ Kit (ZymoResearch, Irvine, California) according to the manufacturer’s instructions. Purified RNAs were submitted to additional DNase digestion in solution and cleaned up with the RNA Clean & Concentrator™-5 Kit (ZymoResearch, Irvine, California).

Total RNAs were reverse transcribed using 0.5 μM oligo (dT) and RevertAid H Minus Reverse Transcriptase (Thermo Scientific, Waltham, USA) according to the manufacturer’s instructions. In the case of IME4 transcript analyses, instead of oligo (dT) the forward and reverse primers ZrIME4F1 and ZrIME4R1 were used for the cDNA synthesis of AS-IME4 lncRNA and S-IME4 mRNA, respectively (S1 Table). Then a common pair of internal primers (ZrIME4_F2/ZrIME4_R2; S1 Table) was used to amplify the same cDNA fragment from each template, if any. Experiments were carried out with three biological replicates and non-reverse transcribed (NRT) controls were performed for each biological replicate. cDNA template (25 ng) was PCR-amplified with primers listed in S1 Table and successful amplification was checked by electrophoresis in 2% agarose gel in 0.5X TBE Buffer.

Quantitative PCR (qPCR) assays

For relative expression level analysis, qPCR reactions were performed using 1 ng/μl of cDNA, 0.3 μM of each primer, and the Maxima™ SYBR Green/ROX qPCR Master Mix (Fermentas, USA) according to the manufacturer’s instructions. Primers were designed to selectively amplify MATα1 exons as well as different copies of MATα1, MATα2 and HO genes (S1 Table). PCR efficiency was in silico predicted for each primer set using the open source tool Primer Efficiency (http://srvgen.upct.es/index.html; [42]). Preliminary, the presence and the estimated size of amplicons were checked by RT-PCR on cDNA from stationary-phase cells (three biological replicates) under standard conditions, whereas unspecific amplifications and primer-dimer formation were checked by melting curve analysis after RT-qPCR assay. All qPCR reactions were run in the Applied Biosystems 7300 Real-Time PCR instrument (Applied Biosystem, Foster City, CA, USA). The Z. rouxii housekeeping gene ACT1 (ZrACT1; XM002497273) was used as a reference gene, according to Leandro et al. [43]. The relative expression of different gene transcripts was calculated by the ΔΔCt method and converted to the relative expression ratio (2−ΔΔCt) for statistical analysis [44]. For this purpose, a dilution series of standard points from pooled control cDNAs was exploited (concentration range 10–0.02 ng/μl) in technical triplicates. Amplification profiles, baselines and thresholds were analysed with 7300 SDS 1.4. PCR-Miner [45] [46] was applied as an alternative method when the low gene expression level did not provide enough dynamic range to build a reliable standard curve. In this case, reaction efficiency and the fractional cycle number at threshold (Ct) were estimated by relying on the kinetics of individual PCR reactions containing 5 ng of cDNA template from three biological replicates (salt-treated and controls), including six technical replicates each.

Results

Morphological characterization and mating test

We first assessed the mating behaviour of the ATCC 42981 strain in pure and mixed cultures with the Z. rouxii mating partners CBS 4837 (matting-type a) and CBS 4838 (matting-type α), respectively. Examination under the microscope did not show any evidence of mating reaction of ATCC 42981 cells neither with Z. rouxii CBS 4837 or CBS 4838 tester strains, even after 3 weeks of incubation both on MEA and 6% NaCl-MEA media (data not shown). Even if rare
mating events cannot be ruled out, this result suggests that the ATCC 42981 strain did not respond to Z. rouxii pheromone signalling or that there was an imbalanced or deficient organization in MTL loci. Based on our previous observations [25], ATCC 42981 cells grown on MEA medium displayed an adhesive phenotype with clamps on mother and daughter cells that remained attached to each other, but they were not able to form ascii in pure culture (data not shown). In contrast, Z. sapae ABT 301\textsuperscript{T} rarely forms ascii at the same conditions [25]. As salt has been reported to induce meiosis in Zygosaccharomyces yeasts [24], we tested the capability of the ATCC 42981 strain to form ascospores in YNB5%G NaCl and 6%NaCl-MEA media. No conjugated ascii were observed over time in any media tested. Furthermore, an increase in the adhesive phenotype was detected in cells grown under salt stress (data not shown).

**Isolation and characterization of MTL\textalpha loci in ATCC 42981**

Three MTL\textalpha loci, termed ZsMTL\textalpha copies 1 to 3, have been previously described in Z. sapae as regions containing phylogenetically distinct \textalpha 1 and \textalpha 2 genes, respectively [32]. To establish whether strain ATCC 42981 also displays a similar genetic configuration, we exploited a PCR walking strategy. This assay relies on MTA\textalpha copy-specific primers designed on the corresponding ZsMTL\textalpha cassettes [32]. PCR reactions were positive using MTL\textalpha copies 1 and 2-specific primers, whereas negative results were scored for any MTL\textalpha copy 3-specific primer set tested. We sequenced two distinct MTL\textalpha loci (hereafter referred to as MTL\textalpha copies 1 and 2) in ATCC 42981. Both loci consisted of two MAT genes encoding \textalpha 1 and \textalpha 2 proteins placed in opposite directions and separated by an intervening region of 343 bp. The MAT\textalpha 1 genes from MTL\textalpha loci 1 and 2 (termed MAT\textalpha 1 copies 1 and 2, respectively) were divergent from each other for 17 nt substitutions, whereas the MAT\textalpha 2 from MTL\textalpha loci 1 and 2 (termed MAT\textalpha 2 copies 1 and 2, respectively) displayed 16 nt substitutions. The deduced proteins MAT\textalpha 1 and MAT\textalpha 2 copies 1 were 87.00% and 78.67% identical to the deduced protein MAT\textalpha 1 and MAT\textalpha 2 copies 2, respectively. BLASTp search against the NCBI database showed that proteins MAT\textalpha 1 and MAT\textalpha 2 copies 1 are more similar to Z. rouxii orthologs (termed ZrMAT\textalpha 1 and ZrMAT\textalpha 2) than the paralogous proteins MAT\textalpha 1 and MAT\textalpha 2 copies 2 in the ATCC 42981 genome.

The NJ-based phylogenetic analysis was carried on MAT\textalpha 1 sequences from representative WGD and non-WGD species. As expected, ATCC 42981 MAT\textalpha 1 copy 2 protein did not group to Z. rouxii MAT\textalpha 1, but was instead clustered separately (bootstrapping value of 99%) together with ZsMAT\textalpha 1 copy 2 (Fig 1). The alignment of ATCC 42981 MAT\textalpha 1 copies with Z. rouxii and S. cerevisiae MAT\textalpha 1 proteins revealed a region of high similarity inside the MAT\textalpha-HMG domain [41] with three predicted conserved \textalpha helices (data not shown).

Phylogeny inferred from the MAT\textalpha 2 aa sequences of WGD and non-WGD species showed a tree topology congruent with the species relationships established using MAT\textalpha 1 sequences. The ATCC 42981 genome harbours two MAT\textalpha 2 variants that are related but phylogenetically distinct because of the high level of amino acid divergence (Fig 2). In particular, ATCC 42981 MAT\textalpha 2 copy 1 clustered with Z. rouxii MAT\textalpha 2 (bootstrapping 100%), whereas MAT\textalpha 2 copy 2 was strictly related to Z. sapae MAT\textalpha 2 copy 2 (bootstrapping 100%). Both copies contained a conserved HD1 class homeodomain, consisting of a three-helix globular domain which contacts both major groove bases and the DNA backbone [47] [48] (data not shown). However, portions of the protein outside the homeodomain, which mediate interactions with accessory proteins, had a different degree of conservation.

**Isolation and characterization of MTL\textalpha loci in ATCC 42981**

In Z. sapae the \textalpha -idiomorph encoding MTL locus harbours a MAT\textalpha 1-coding ORF (ZsMAT\textalpha 1) identical to the Z. rouxii orthologue (ZrMAT\textalpha 1) and a MAT\textalpha 2-coding ORF (ZsMAT\textalpha 2), which
showed a 26-bp deletion compared to ZrMATα2, resulting in a 9-amino acid shorter protein [32]. To identify MTLα loci in the ATCC 42981 genome, a PCR strategy was used based on primers designed on the MATα1 and MATα2 gene sequences in Z. rouxii and Z. sapae. We found three MTLα loci, each harbouring MATα1 and MATα2 genes in opposite directions separated by a 279-bp long intergenic sequence, which differed for a single SNP (T/G) compared to the corresponding region in Z. sapae.

Sequence alignments and BLAST-type search revealed that all ATCC 42981 MTLα loci displayed identical MATα1 genes (100% identity to ZrMATα1 and ZsMATα1), whereas differed from each other in the MATα2-coding ORFs. One MATα2-coding gene (referred to as the
Fig 2. Phylogenetic analysis of MATα2 proteins. The neighbour-joining (NJ) tree shows the phylogenetic relationships between the allodiploid strain ATCC 42981 and other hemiascomycetes inferred from MATα2 proteins. Numbers on branches indicate bootstrap support percentages (1,000 pseudoreplicates) higher than 60% from NJ. The red branch indicates Z. rouxii complex, which includes Z. rouxii MATα2, ATCC 42981 MATα2 copies 1 and 2, and Z. sapae MATα2 copies 1 to 3 sequences. The dark dot indicates WGD species, whereas the dark triangle indicates non-WGD species with the HO gene.

doi:10.1371/journal.pone.0152558.g002
MATa2 copy 1) was 100% identical to *Z. rouxii* counterpart. Another MATa2 gene (termed the MATa2 copy 2) encoded an a2 protein, which diverged from ZrMATa2 mainly at the C-terminal end (93.73% identity). A third MATa2 coding sequence, namely MATa2 copy 3, encoded a protein 94.41% and 94.24% identical to ZrMATa2 and ZsMATa2, respectively.

The NJ-tree confirmed the high degree of evolutionary conservation of MATa1 proteins within the *Z. rouxii* complex (bootstrapping 100%) (data not shown). Furthermore, a search with the program Pfam revealed that ATCC 42981 MATa1 proteins conserve the HD2 class homeodomain [48] [49] (data not shown).

Phylogeny inferred from MATa2 amino acid sequences showed that ATCC 42981 MATa2 variants are phylogenetically distinct (Fig 3). In particular, MATa2 copy 1 clustered with *Z. rouxii* and *Z. sapae* MATa2 (bootstrapping 100%), whereas MATa2 copies 2 and 3 clustered together and were distinct from both ZrMATa2 and ZsMATa2. MATa2 proteins from ATCC 42981 conserved a single MATA-HMG box, a class I member of the HMG-box superfamily of DNA-binding proteins (Fig 4), coding a sequence spanning across the Ya and X regions.

**Characterization of *HO* genes**

Using a PCR walking approach, we identified two 1797-bp full-length ORFs in the ATCC 42981 genome, termed *HO* copies 1 and 2, which diverged for 177 transitional and 49 transversional mismatches. The predicted proteins HO copies 1 and 2 shared 100% and 92.47% identity with *Z. rouxii* HO, while they were 100% and 99.83% similar to *Z. sapae* HO copies 1 and 2, respectively.

---

**Fig 3. Phylogenetic analysis of MATa2 proteins.** The neighbour-joining (NJ) tree shows the phylogenetic relationships between the allodiploid strain ATCC 42981 and other hemiascomycetes inferred from MATa2 proteins. Numbers on branches indicate bootstrap support percentages (1,000 pseudoreplicates) higher than 60% from NJ. The red branch indicates the *Z. rouxii* complex, which includes *Z. rouxii* MATa2, *Z. sapae* MATa2 and ATCC 42981 MATa2 copies 1 to 3 sequences. The dark dot indicates WGD species, whereas the dark triangle indicates non-WGD species with the *HO* gene.

doi:10.1371/journal.pone.0152558.g003
NJ-based phylogeny inferred from amino acid HO sequences confirmed that HO copy 2 from ATCC 42981 branched with Z. sapae HO copy 2 with a high level of support (bootstrap- ping 100%), while ATCC 42981 HO copy 1, ZsHO copy 1 and ZrHO clustered together into a separate clade (Fig 5). These results showed that ATCC 42981 HO amino acid sequences are more divergent from each other than from the putative orthologs in Z. sapae. Amino acid alignment of both ATCC 42981 HOs with S. cerevisiae PI-SceI (GenBank: CAA98762), S. cerevisiae HO and Z. sapae HO copies 1 and 2 showed the highest homology in conserved motifs characteristic of intein-encoded LAGLIDADG endonucleases [51–53] (data not shown).

PFGE-Southern blotting showed that both HO copies of ATCC 42981 strain are on chromosome F (1.6 Mbp) (S4 Fig). This chromosomal assignment resembles that of CBS 732T, but differs from that previously found in Z. sapae ABT301T [32]. We assumed that ATCC 42981 HO copies could be placed on the same chromosome or, alternatively, on two PFGE-co-migrating homeologous chromosomes.

Subgenome assignment of mating-type and HO genes

In order to infer the putative parental contributions to the ATCC 42981 MTLα, MATα and HO gene sets, we screened strain NCYC 3042 for all variants of MATα, MATα, and HO genes found in ATCC 42981. According to PCR profiles, successful amplification was obtained for copy 2 sets of MATα1, MATα2 and HO genes in the NCYC 3042 strain, while negative results were obtained for MATα and HO copy 1 genes (S2 Table). PCR results were consistent with the high degree of similarity between the copy 1 set of MATα and HO genes and Z. rouxii orthologs in the CBS 732T genome project (Figs 1, 2 and 5) and suggested that MATα and HO copies 2 originated from a Z. pseudorouxii parental genome. Interestingly, MATα2 copy-
specific screening showed that the NCYC 3042 strain does not possess any $MAT_a^2$ copies found in ATCC 42981 genome (S2 Table).

Reconstruction of the three-cassette system

To assign the $MTL_a$ and $MTL_\alpha$ loci as $MAT$ expression loci, $HMR$ or $HML$ silent cassettes (Fig 6), we exploited direct PCR on $MTL$ flanking regions and long-range PCRs that spanned left and right flanking regions of $MTL$ loci (S2 Fig). Both approaches demonstrated that ZYR0F18524g (termed $CHA_1\_L$ due to its similarity to $CHA_1$ gene) and ZYRO0F18634g ORFs are located at the 5' and 3' ends of the $MTL_\alpha$ copy 1 locus, respectively. Because $HML_\alpha$ silent cassettes are commonly downstream the $CHA_1\_L$ gene [31], we inferred the existence of a $HML_\alpha$ silent cassette containing $MAT_\alpha^1$ and $MAT_\alpha^2$ copies 1 genes (referred to as $HML_\alpha$ copy1) (Fig 6). Long-range PCR products obtained with the primer pairs 2/A and 3/A were positively screened through $MTL_\alpha$ copy 2-specific primers (S2 Fig). Sequencing showed that there are two $MTL_\alpha$ copy 2 loci flanked by the $SLA_2$ gene at the 3' end and by either the $DIC_1$ gene or $CHA_1\_L$ at the 5' end, respectively. It has been reported that the $MAT_\alpha$ expression loci retain the particular gene order $DIC_1\cdot MAT_\alpha\cdot SLA_2$ in the majority of non-WGD species [54]. A similar synteny was found for $MTL_\alpha$ copy 2, suggesting that ATCC 42981 possesses an actively
transcribed \( \text{MAT} \alpha \) copy 2 locus (Fig 6). Furthermore, \( \text{CHA1L} \) upstream to the \( \text{MTL} \alpha \) copy 2 supports the presence of an additional \( \text{HML} \alpha \) copy 2 cassette. Finally, we detected the following syntenic orders for \( \text{MTL} \alpha \) loci: DIC1-MTLα copy 1-ZYRO0C18392g, \( \text{CHA1-MLT} \) copy 2-SLA2, \( \text{CHA1-MLT} \) copy 2-ZYRO00C18392g and \( \text{CHA1L-MLT} \) copy 3-ZYRO00C18392g (Fig 6). The gene organization of the \( \text{CHA1-MTL} \) copy 2-SLA2 was syntenic with that of the \( \text{MAT} \alpha \) expression locus in \( \text{Z. rouxii} \) CBS 732T, whereas the gene orders \( \text{CHA1L-MTL} \) copy 2-ZYRO00C18392g and DIC1-MLTa copy 1-ZYRO0C18392g resembled those found in \( \text{HMR} \alpha \) silent loci of haploid \( \text{Z. rouxii} \) strains [31]. In particular, \( \text{CHA1-MLT} \) copy 2-ZYRO00C18392g resembled the \( \text{HMR} \alpha \) cassette found in \( \alpha \)- and \( \alpha \)-idiomorph mixed culture of strain NBRC 1130. Although the synteny \( \text{CHA1L-MTL} \) copy 3-ZYRO0C18392g has been never found in \( \text{Z. rouxii} \) strains, \( \text{CHA1L} \) and ZYRO00C18392g ORFs generally surround silent donor cassettes. Overall, these evidences support the notion that ATCC 42981 possesses one \( \text{MAT} \alpha \) copy 2 expression locus and three \( \text{HMR} \alpha \) cassettes, referred to as \( \text{HMR} \alpha \) copies 1 to 3 (Fig 6).

The chromosomal arrangement of \( \text{MAT}, \text{HMR} \) and \( \text{HML} \) cassettes was established by Southern blot analysis on PFGE-separated chromosome bands. Although PFGE-Southern blotting failed to clearly resolve the highest molecular weight chromosomes E, F and G.

Fig 6. Inferred genomic organization around MAT-like loci in the ATCC 42981 allodiploid genome. Two sex homologous/homeologous chromosome pairs are depicted, namely F/F’ and G/G’. Chromosome G bears the \( \text{MAT} \alpha \) copy 2 expression locus, which is linked to the putative silent cassettes \( \text{HML} \alpha \) copy 2, whereas chromosome F’ bears the \( \text{MAT} \alpha \) copy 2 expression locus, which is linked to the putative silent cassette \( \text{HML} \alpha \) copy 1. Chromosomes F and G’ harbour three \( \alpha \)-idiomorph \( \text{HMR} \) loci, which differ for \( \text{MAT} \alpha \) genes. The \( \text{HMR} \) copy 2 locus is on chromosome F, while the \( \text{HMR} \) copies 1 and 3 loci are on chromosome G’. Blue arrows represent \( \text{Z. rouxii} \)-like (blue bordered) and \( \text{Z. sapae} \)-like (light green bordered) \( \alpha \)-idiomorph loci. Red arrows indicate \( \alpha \)-idiomorph loci with different \( \text{MAT} \alpha \) genes: copy 1 (dark red surrounded), copy 2 (dark green), copy 3 (grey), respectively. Chromosomal organization of the three-cassette system in \( \text{Z. rouxii} \) haploid strain CBS 732T was reported for comparative purposes according to Souciet et al. [34]. \( \text{CHA1L} \) indicates the ZYRO0F18524g locus, while dark green and light blue arrows indicate the ZYRO0F18634g and ZYRO0C18392g loci, respectively at the right side of mating-type loci.

doi:10.1371/journal.pone.0152558.g006
spanning from 1.6 to 2.2 Mbp (S5 Fig), hybridization of PFGE-Southern blot with a MATα1-specific probe (suitable to recognize both MATα1 copies 1 and 2) resulted in a double band spanning from chromosome F to G for Z. rouxii CBS 732T and strain ATCC 42981 (S5 Fig). However, differently from CBS 732T, in strain ATCC 42981, the MATα1-specific probe bound less chromosome F than chromosome G, leading to two bands with different signal intensities. This result suggests that chromosome G harbours more copies of the MTLα cassette than chromosome F. When the same analysis was carried out with a MATα1-specific probe, we obtained a double band spanning from chromosome F and G of ATCC 42981 karyotype (S5 Fig). These studies suggest that MTLα loci reside on at least two of the chromosomes spanning from 1.6 and 2.2 kbp. According to PFGE-Southern blotting results, the ATCC 42981 strain arranges two expressed MAT loci (MATα and MATα copies 2), three HMRα and two HMLα silent cassettes on two homeologous chromosome pairs (G/G' and F/F'). Taking into consideration that linked MAT and HML loci are located on a different chromosome compared to HMR [31] [54], we infer that the scenario depicted in Fig 6. Chromosome G putatively contains the expression locus MATα copy 2, which is linked to the HMLα copy 2 cassette, while chromosome F hosts the HMRα copy 2 silent cassette. The expression locus MATα copy 2 is putatively located on chromosome F', linked with HMLα cassette copy 1. Chromosome G' provides the two remaining silent HMRα copies 2 and 3 cassettes (Fig 6).

Gene expression analysis

Non-quantitative RT-PCR confirmed that the expression loci MATα and MATα copies 2 are actively transcribed in ATCC 42981 cells grown in standard conditions (Fig 7). Congruently to the cassette reconstruction, the MATα1 and MATα2 genes from HMLα copy 1 were silent. According to the regulatory circuit reported in S. cerevisiae diploid cells, we expected that h-sg sets are silenced by a functional a1-α2 heterodimer. In contrast, gene expression profile analysis revealed that the MATα1 and HO genes were actively transcribed in the ATCC 42981 allo-diploid strain. Similarly in haploid CBS 732T unstressed cells HO gene was actively transcribed (data not shown).

To verify whether other members of the h-sg set were expressed, we tested the presence of S-IME4 and AS-IME4 transcripts. RT-PCR analysis showed that salt-stressed ATCC 42981 diploid cells only transcribed AS-IME4, while S-IME4 was not detected (Fig 7). This pattern is similar to that displayed by haploid cells of the reference strain CBS 732T, which does not show any PCR products for S-IME4 specific RT-PCR (data not shown). The identity of AS-IME4 PCR products was confirmed by direct sequencing. We concluded that ATCC 42981 is a diploid strain that only transcribes the AS-IME4 transcript.

Transcriptional response to hyperosmotic stress

After hyperosmotic stress, ATCC 42981 cells exhibit a different expression profile for both mating-type and HO genes compared to controls. Salt stress induced higher transcript levels of MATα1 and MATα1 copy 2 (6.2- and 9.3-fold, respectively). On the contrary, MATα2 copy 2 expression was lower than that in controls (2.4-fold). HO copy 1 was slightly down-regulated (2.8-fold), whereas HO copy 2 was up-regulated (more than 5.3-fold) compared to controls (Fig 8).

Discussion

Previous studies demonstrated that outcrossing is a very rare event (approximately once every 50,000/110,000 generations) in yeasts as S. cerevisiae [55], Saccharomyces paradoxus [56] and Lachancea kluyveri [57]. Nevertheless, it represents a potential source of phenotypic variability
**Fig 7. Expression pattern of mating-type, HO and IME4 genes.** Panel A reports positive amplified cDNAs generated with MAT and HO copy variant-specific primers from ATCC 42981 unstressed cells in stationary growth phase. Panel B depicts IME4-specific PCR products generated from CBS 732T and ATCC 42981 cDNA IME4 antisense (AS-IME4 lncRNA) and sense transcripts (S-IME4 mRNA), respectively. +/- RT indicates addition of reverse transcriptase to the cDNA synthesis reaction. For each RT-PCR reaction gDNA was used as positive control. Abbreviations: AS, anti-sense long non-coding RNA; S, sense mRNA.

doi:10.1371/journal.pone.0152558.g007

**Fig 8. Differential expression by quantitative real-time PCR of mating-type and HO genes in ATCC 42981 cells after hyperosmotic stress.**

Expression of target genes was normalized on the reference ZrACT1 (GenBank: XM002497273). Fold change was measured by ΔΔCt or PCR Miner methods and reported as the mean (± SEM) of three biological replicates. * indicates significant difference from controls as measured by independent Student's t-tests (*P<0.05, **P<0.01).

doi:10.1371/journal.pone.0152558.g008
Chimeric Sex System Determines and Cell-Type Identity Dysregulation

(novelty) and has been extensively exploited for the genetic improvement of microbial cell factories [58]. Although many inter-mating allopatric yeasts give rise to viable hybrids, postzygotic isolating barriers [39] prevent gene flow between species and contribute to the process of speciation [60]. Complete sets of orthologous genes are expected in yeast hybrids immediately after the merging of two parental genomes. These patterns can undergo extensive homogenization processes over evolutionary time through intragenic recombination, gene conversion and differential gene loss, shaping the offspring’s genome architectures [61] [62]. Gordon and Wolfe [29] did not find any traces of gene losses in the ATCC 42981 strain, arguing that the allopolyploidization was so recent that its genome has not had enough time to decay. These authors hypothesized that one parental subgenome resembles Z. rouxii CBS 732T and the other Z. pseudorouxii (nom. inval.) NCYC 3042. However, ATCC 42981 karyotype cannot be simply considered as an additive result between the putative parental counterparts, suggesting that some structural rearrangements have occurred in the ATCC 42981 genome compared to Z. rouxii and Z. pseudorouxii [28] [29]. Similarly, the Z. rouxii CBS 732T strain contains a SOD2-22 gene variant [63], while Z. pseudorouxii NCYC 3042 has SOD2 [30]. The detection of SOD2 and SOD22 genes, but not the SOD2-22 variant in the ATCC 42981 genome, further confirms that the subgenome complements are similar but not identical to those found in Z. rouxii CBS 732T and in Z. pseudorouxii NCYC 3042.

In an allodiploid karyotype, two homeologous sex chromosomes are expected, bearing a and α-idiomorph expression loci, respectively. Previous works demonstrated that in haploid Z. rouxii and diploid Z. sapae, HMR cassettes are located on a different chromosome compared to the MAT-HML linkage, accounting for the presence of two sex chromosomes [31] [32]. Therefore, in a recent allodiploid Zygosaccharomyces genome, we expected to identify two sex chromosome pairs. Accordingly, the ATCC 42981 strain displays an α/α genotype provided by the contribution of two sex chromosome pairs. The first sex chromosome pair consists of chromosomes G (MATα copy 2–HMLα copy 2) and F (HMRα copy 2). This subgenome complement only partially resembles that found in Z. pseudorouxii NCYC 3042 strain [30] which might have contributed to this complement with MTLα copy 2. If Z. rouxii contributed to the second chromosome set, we would expect to detect a Z. rouxii MATα2 copy 1 gene at the expression locus. In contrast, chromosome F (carrying MATα copy 2–HMLα copy 1) and G (carrying HMRα copy 1–HMRα copy 3) probably derived from a rearranged Z. rouxii genome complement [29], consistently with the mating-type loci being the hotspot of ectopic recombination in Z. rouxii strains [31]. Furthermore, an extra HMRα copy 3 was detected harbouring another partially divergent MATα2 ORF, which might have derived from another non-Z. rouxii parental strain or resulted from an ectopic recombination event in the allodiploid ancestor. All these findings are not consistent with a simple additive set of mating-type loci recently assembled in the allodiploid genome and indicate that ATCC 42981 T- subgenome differs from its putative Z. rouxii CBS 732T counterpart. Watanabe et al. [31] found that the copy 2 variant of MATα2 gene is at the HMR silent cassette of strain NBRC1130 which undergone mating-type switching, whereas copy 1 variant occurs in MAT expression locus. This genome assortment is speculative to that found in ATCC 42981 T-subgenome and hints that different MATα2 copies co-exist within Z. rouxii species, harboured by MTL loci flanked by different genes at 5’ end. We speculated that these copy variants could be relicts of past mating-type switching events. Interestingly, Z. rouxii possesses a hybrid regulation system targetting a-specific genes (a-sgs), which consists of a2-mediated activation and a2-mediated repression [64]. In haploid cells, MATα2 proteins with diverging C-terminal portions could be functionally equivalent in mediating the activation of the a-sg program.

Beyond divergent mating-type loci, the ATCC 42981 strain has also been reported to possess two divergent HO genes. Orthologous genes coming from each of the parental species
should appear as paralogs in standard analyses because they are homologous genes encoded by the same genome [65]. However, the degree of divergence identified between *HO* gene copies 1 and 2 excludes their origin as paralogs from a recent gene duplication event and hints that *HO* copy 2 originated from the *Z. pseudorouxiyi* parental genome. The pattern of neutral mutations detected in the endonuclease functional domains indicates that both *HO* genes have been exposed to the same selective pressure.

Gene regulatory networks, such as cell-type specification circuit, have been shown to evolve significantly over time [66]. When the divergent components of these circuits are forced to interact in a hybrid background, positive or antagonistic epistatic interactions may take place [67]. Among the mechanisms cited to explain the observed loss of hybrid fertility, the Bateson-Dobzhansky-Muller (BDM) model proposes that hybrid sterility results from the lack of interaction or the malfunctioning of interacting alleles derived from divergent genomes [3] [68–70]. However, in some cases, inter-specific protein assemblies have been reported to generate novelties in protein-protein networks untested by selection in hybrid species [71] [72]. The reconstruction of the sex chromosome architecture in the ATCC 42981 genome indicates that the transcriptional regulators encoded at the *MAT* expression loci are phylogenetically divergent.

In diploid cells, a1–α2 heterodimer acts as a master regulator of the shift from mitosis to meiotic growth under appropriate meiosis-inducing stimuli, as well as a repressor of the h-sgs under standard growth conditions. H-sgs contain the *MATα1* gene, which the a1–α2 heterodimer should repress by binding the bidirectional promoter located between *MATα1* and *MATα2* ORFs [12]. In allo-diploid ATCC 42981, the lack of *MATα1* silencing could be due to: the lack of a1 and α2 proteins; the failure in the heterodimer formations; the inability of putative chimeric a1 and α2 subunits to positively interact with *cis*-regulatory promoter sequences. Similarly to ATCC 42981, *S. cerevisiae* diploids showing mutations in either a1 or α2 transcriptional factors fail to turn off *MATα1* and other h-sgs [73] [74]. Additionally, Strathern et al. [75] reported that a diploid mutant with a truncated *MATα2* defective allele displays a weak α phenotype and is unable to sporulate. Overall, these evidences suggest that in ATCC 42981 the cell-type specification circuit is ineffective in repressing the *MATα1* gene possibly due to, among other factors, a no functional chimeric a1–α2 heterodimer.

In diploids, the entry into meiosis requires the inhibition of InC RNA AS-IME4 by a functional a1–α2 heterodimer [21]. Conversely, our results show that the ATCC 42981 strain produces anti-sense transcripts for the *IME4* gene, leading to clonality as its unique mode of reproduction. The haploid-like transcriptional pattern displayed by stressed ATCC 42981 cells could also account for its increase of adhesive phenotype. In haploids grown under stress cues, a complex network of signalling modules and transcriptional factors induce clamp formation and pseudohyphal growth in order to enhance mating efficiency and chance of survival [76]. The *FLO11* gene has been reported to be a key determinant of the adhesion phenotype under the positive control of the RME1 transcriptional factor. In diploids, the *RME1* gene is silenced by a functional a1–α2 heterodimer, leading to the entry into meiosis in response to environmental cues. Because we observed more clamps in salt-stressed compared to unstressed ATCC 42981 cells, we argue that the chimeric a1–α2 heterodimer is also partially ineffective in repressing *RME1* transcription.

In *S. cerevisiae*, a redundant regulatory network accounts for the three-level regulation of mating-type interconversion catalysed by the *HO* endonuclease, namely cell-type control (*HO* gene is expressed in a or α haploid cells); mother–daughter control (*HO* is transcribed in the mother but not in the daughter cells); and cell-cycle control (*HO* is expressed during the late G1 phase of the cell cycle after the point of commitment to the next cell cycle) [77]. Three transcriptional repressors are involved: the a1–α2 heterodimer, *SIN3* and *SIN6*, respectively. These three types of negative constraints must be relieved in order for the *HO* gene to be transcribed.
Allodiploid ATCC 42981 cells express HO mRNAs, probably due to the lack of these types of controls or to partial incompatibilities among transcriptional factors. However, the presence of HO transcripts does not imply that ATCC 42981 undergoes mating-type interconversion. Indeed, the transcriptional analysis of MAT expression loci showed that only the expected MATα copy 2 and MATα copy 2 transcripts are detected in salt-stressed cells at the stationary phase. The failure of mating-type switching induced by salt stimuli could be due to either a HO post-transcriptional control or to the lack of a fully functional network controlling the DSB-initiated gene conversion. However, other effectors could be responsible for this event, as in haploid Z. rouxii strains HO deletion determines only a slight decrease in mating-type switching frequency [31]. The Z. rouxii species complex emerged after the ancestor of hemiascomycetous yeasts had diverged from other families, such as Kluyveromyces. K. lactis has a non-functional copy of the HO gene [78] and performs mating-type switching by an alternative transposase-mediated mechanism [79]. As it is the first non-WGD clade with a functional HO gene, the Z. rouxii complex is likely to retain remnants of both mechanisms.

In haploid Z. rouxii, mating and the subsequent zygote formation occur immediately before sporulation mainly under salt stress [24][80]. Therefore, we investigated how the ATCC 42981 cells modulate the transcription of genes coding the α1 and α2 subunits, as well as that of their downstream h-sg targets MATα1 and HO in response to long-term hypersaline stress. In diploid cells, the silencing of h-sg MATα1 and HO and the positive regulation of the MATα2 gene are controlled by a working α1-α2 heterodimer. The inability of the ATCC 42981 chimeric α1-α2 heterodimer to bind to the h-sg promoter regions may account for the observed up-regulation of MATα1 and HO genes. Functional defects of the chimeric α1-α2 heterodimer could be due to gene incompatibility between two divergent subunits and/or to the transcriptional imbalance of their encoding genes. ATCC 42981 cells could attempt to overcome these functional deficiencies by over-expressing the components of the α1-α2 transcriptional factor. Con- gruently, we observed an up-regulation of the α1 transcript. A similar up-regulation of the other heterodimer subunit α2 should be expected. In contrast, MATα2 gene expression appears to be down-regulated, even at a small level, suggesting an imbalance in the co-regulation of α1 and α2 subunits.

In conclusion, we demonstrated that allodiploid ATCC 42981 cells display a MATα/MATα genotype with a chimeric sex-determination system originating from the co-existence of two different parental genome complements. The protein-protein interaction incompatibility between divergent α1 and α2 subunits could switch-off the meiosis commitment genes, contributing to ATCC 42981 allodiploid sterility. The presence of a chimeric α1-α2 heterodimer promotes an unusual haploid-like transcriptional profile in cells recovered at the stationary phase and after exposure to meiosis-inducing stimuli. To the best of our knowledge, this is the first clue that the BDM interaction between the divergent α1 and α2 subunits may act as a bottleneck, preventing genetic exchanges among Zygosaccharomyces species. Recently, a novel scenario has been proposed for yeast evolution, where two ancient non-WGD ancestral species have given rise to an allodiploid cell that has doubled its genome in order to restore fertility (with a possible interval of many mitotic generations between these two events) [81]. Interestingly, one of the possible parents exhibits phylogenetic affinities with the non-WGD Z. rouxii clade. However, there are several critical open questions that still need to be answered, such as: how the genome duplication event took place and how the mechanism of restoring fertility operated [65]. Allodiploid ATCC 42981 and other strains belonging to the Zygosaccharomyces mosaic lineage [23] could serve as promising models to shed light on the transcriptional network incompatibility underlying hybrid sterility at an incipient stage of speciation, and more in general, on yeast genome evolution.
Supporting Information

S1 Fig. Outlined experimental strategy used for characterizing \textit{MTL}α (A) and \textit{MTLa} loci (B) in the ATCC 42981 genome. Panel A shows Zs\textit{MTLa} variants represented in blue and surrounded in grey (copy 1), green (copy 2) and orange (copy 3), respectively, while panel B represents \textit{MTLa} loci coloured in red. Primer pairs specific for Zs\textit{MTLa} copies 1 to 3 are arbitrarily referred to as α\textit{Cp1-P}, α\textit{Cp2-P}, and α\textit{Cp3-P}. Zs\textit{MATa1} and Zs\textit{MATa2}-specific primer pairs are arbitrarily referred to as Pa1 and Pa2, while the primer pair termed Pa12 spans the complete \textit{MATa2} gene. Solid grey arrows indicate generic flanking genes, dotted borders represent uncompleted sequences and small arrows (solid) designate gene-specific primers. Primer sequences are reported in S1 Table, according to [31] and [32]. Abbreviations: Zs, \textit{Zygosaccharomyces sapae}; cp, copy; P, primer; for, forward; rev, reverse. (TIF)

S2 Fig. Polymerase chain reaction (PCR)-based strategies used for determining the system of cassette-based arrangement in the ATCC 42981 genome. Forward and reverse \textit{MTL}-specific internal primers were used to screen PCR products obtained using all possible combinations of primers spanning putative \textit{MTL}-flanking genes (semi-nested PCR approach); in cases of negative results, 5' and 3' PCR walking was performed using all possible combinations of \textit{MTL}-specific internal primers and \textit{MTL}-flanking gene primers (direct PCR approach). Small arrows (solid) indicate gene-specific primers and degenerate primers (dotted lines). CHAI\textsubscript{L} indicates the ZYR00F18524g locus. Primer sequences are reported in S1 Table, according to [31] and [32]. Abbreviations: cp, copy; for, forward; rev, reverse. (TIF)

S3 Fig. Outlined experimental approach used for \textit{HO} gene characterization in the ATCC 42981 genome. Dotted lines represent undetermined sequences. Primer sequences are reported in S1 Table. Abbreviation: Zs\textit{HO}, \textit{Zygosaccharomyces sapae} \textit{HO} gene. (TIF)

S4 Fig. Chromosomal mapping of ATCC 42981 \textit{HO} genes. Chromosomes were separated by PFGE for ATCC 42981 (1), \textit{Z. rouxii} CBS 732\textsuperscript{T} (2), and \textit{Z. sapae} ABT301\textsuperscript{T} (3). Southern blotting analysis was carried out with probe labelled to \textit{HO} genes. M indicates the chromosomal size ladder (\textit{Saccharomyces cerevisiae} S288C, BioRad Laboratories) in megabase pairs. ATCC 42981 chromosomes are indicated in uppercase letters (from A to G). (TIF)

S5 Fig. Chromosomal mapping of ATCC 42981 \textit{MTL}α and \textit{MTLa} loci. Chromosomes were separated by PFGE for \textit{Z. rouxii} CBS 732\textsuperscript{T} (1) and ATCC 42981 (2). Southern blotting analyses were carried out with probes labelling the α- and α-idiomorph loci. The left panel shows signals from \textit{MTL}α loci \textit{Zygosaccharomyces rouxii} CBS 732\textsuperscript{T} (1) and the ATCC 42981 strain (2), respectively. The right panel reports separated chromosomes and signals from \textit{MTLa} loci in ATCC 42981 (2). M indicates the chromosomal size ladder (\textit{Saccharomyces cerevisiae} S288C, BioRad Laboratories) in megabase pairs. ATCC 42981 chromosomes are indicated in uppercase letters (from A to G). (TIF)

S1 Table. List of primers used in the present study. 
(XLS)

S2 Table. PCR-based subgenome assignment of mating-type and \textit{HO} gene copies in \textit{Zygosaccharomyces pseudorouxii} (nom. inval.) NCYC 3042. 
(DOCX)
**Author Contributions**

Conceived and designed the experiments: LS SC. Performed the experiments: LS MB SC. Analyzed the data: LS MB SC. Contributed reagents/materials/analysis tools: SC PG. Wrote the paper: LS SC.

**References**

1. Van de Peer Y, Maere S, Meyer A. The evolutionary significance of ancient genome duplications. Nat Rev Genet. 2009; 10(10): 725–32. doi:10.1038/nrg2600 PMID: 19652647

2. Kellis M, Patterson N, Endrizzi M, Birren B, Lander ES. Sequencing and comparison of yeast species to identify genes and regulatory elements. Nature. 2003; 423(6937): 241–54. PMID: 12748633

3. Scannell DR, Butler G, Wolfe KH. Yeast genome evolution-the origin of the species. Yeast. 2007; 24(11): 929–42. org/doi:10.1002/yea.1515 PMID: 17907807

4. Johnson NA. Hybrid incompatibility genes: remnants of a genomic battlefield? Trends Genet. 2010; 26(7): 317–25. doi:10.1016/j.tig.2010.04.005 PMID: 20621759

5. Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T, et al. Major ecological transitions in wild sunflowers facilitated by hybridization. Science. 2003; 301(5637): 1211–6. PMID: 12907807

6. Alipaz JA, Karr TL, Wu CI. Evolution of sexual isolation in laboratory populations: fitness differences between mating-types and the associated hybrid incompatibilities. Am Nat. 2005 Apr 1; 165(4): 429–38. PMID: 15791535

7. Otto SP, Gerstein AC. The evolution of haploidy and diploidy. Curr Biol. 2008; 18(24): 1121–8. doi:10.1016/j.cub.2008.11.021 PMID: 19064385

8. Nolte AW, Tautz D. Understanding the onset of hybrid speciation. Trends Genet. 2010; 26(2): 54–8. doi: 10.1016/j.tig.2009.12.001 PMID: 20044166

9. Zörgö E, Chwialkowska K, Gjuvsland AB, Garré E, Sunnerhagen P, Liti G, et al. Ancient evolutionary trade-offs between yeast ploidy states. PLoS Genet. 2013; 9(3): e1003388. doi:10.1371/journal.pgen.1003388 PMID: 23555297

10. Heitman J. Microbial genetics: Love the one you’re with. Nature. 2009; 460(7257): 807–8. doi:10.1038/nature08011 PMID: 19675638

11. Knop M. Evolution of the hemiascomycete yeasts: On life styles and the importance of inbreeding. BioEssays. 2006; 28(7): 696–708. PMID: 16929561

12. Haber JE. Mating-type genes and MAT switching in Saccharomyces cerevisiae. Genetics. 2012; 191(1): 33–64. doi:10.1534/genetics.111.134577 PMID: 22555442

13. Granek JA., Kayıkçı Ö, Magwene PM. Pleiotropic signaling pathways orchestrate yeast development. Curr Opin Microbiol. 2011; 14(6): 676–81. doi: 10.1016/j.mib.2011.09.004 PMID: 21962291

14. Goutte C, Johnson AD. a1 protein alters the DNA binding of alpha2 repressor specificity. Cell. 1988; 52:875–82. PMID: 3127056

15. Herskowitz IR. Life cycle of the budding yeast Saccharomyces cerevisiae. Microbiol Rev. 1988; 52(4): 536–53. PMID: 3070323

16. Mathias JR, Hanlon SE, O’Flanagan RA, Sengupta AM, Vershon AK. Repression of the yeast HO gene by the MATa2 and MATa1 homeodomain proteins. Nucleic Acids Res. 2004; 32(22): 6469–78. PMID: 15598821

17. Covitz PA, Herskowitz I, Mitchell AP. The yeast RME1 gene encodes a putative zinc finger protein that is directly repressed by a1–α2. Genes Dev. 1991; 5(11): 1982–9. PMID: 19369899

18. Mitchell AP, Herskowitz I. Activation of meiosis and sporulation by repression of the RME1 product in yeast. Nature. 1986 Feb 27; 319(6056): 738–42. PMID: 3513021

19. Van Dyk D, Hansson G, Pretorius IS, Bauer FF. Cellular differentiation in response to nutrient availability: the repressor of meiosis, Rim1p, positively regulates invasive growth in Saccharomyces cerevisiae. Genetics. 2003; 165(3): 1045–58. PMID: 14668363

20. Shah JC, Ciancy MJ. IME4, a gene that mediates MAT and nutritional control of meiosis in Saccharomyces cerevisiae. Mol Cell Biol. 1992; 12(3): 1078–86. PMID: 1545790

21. Hongay CF, Grisafi PL, Galitski T, Fink GR. Antisense transcription controls cell fate in Saccharomyces cerevisiae. Cell. 2006; 127(4): 735–45. PMID: 17110333

22. Gelfand B, Mead J, Bruning A, Apostolopoulos N, Tadigotla V, Nagaraj V, et al. Regulated antisense transcription controls expression of cell-type-specific genes in yeast. Mol Cell Biol. 2011; 31(8): 1701–9. doi: 10.1128/MCB.01071-10 PMID: 21300780
23. Solieri L, Chand Dakal T, Croce MA, Giudici P. Unravelling genomic diversity of *Zygosaccharomyces rouxii* complex with a link to its life cycle. FEMS Yeast Res. 2013; 13(3): 245–58. doi: 10.1111/1567-1364.12027 PMID: 23279556

24. Mori H. Life cycle in a heterothallic haploid yeast, *Saccharomyces rouxii*. J Ferment Technol. 1973; 51: 379–392.

25. Solieri L, Dakal TC, Giudici P. *Zygosaccharomyces sapae* sp. nov., isolated from Italian traditional bal-samic vinegar. Int J Syst Evol Microbiol. 2013; 63(Pt 1): 364–71. doi: 10.1099/ijs.0.043323-0 PMID: 23024146

26. Solieri L, Cassanelli S, Croce MA, Giudici P. Genome size and ploidy level: new insights for elucidating relationships in *Zygosaccharomyces* species. Fungal Genet Biol. 2008; 45(12): 1582–90. doi: 10.1016/j.fgb.2008.10.001 PMID: 18952188

27. Solieri L, Dakal TC, Bicciato S. Quantitative phenotypic analysis of multistress response in *Zygosaccharomyces rouxii* complex. FEMS Yeast Res. 2014; 14(4): 586–600. doi: 10.1111/1567-1364.12146 PMID: 24533625

28. Pribylova L, De Montigny J, Sychrova H. Osmoresistant yeast *Zygosaccharomyces rouxii*: the two most studied wild-type strains (ATCC 2623 and ATCC 42981) differ in osmotolerance and glycerol metabolism. Yeast. 2007; 24(3): 171–80. PMID: 17351908

29. Gordon JL, Wolfe KH. Evolution of the MAT locus and its Ho endonuclease in yeast species. Proc Natl Acad Sci U S A. 2004; 101(6): 1632–7. PMID: 14745027

30. Larkin MA, Blackshields G, Brown NP, Chenna R, Mcgettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23(21): 2947–8. PMID: 17846036

31. Altschul S, Madden T, Schaffer A, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids Res. 1997; 25(17): 3389–405. PMID: 9254994

32. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013; 30(12): 2725–9. doi: 10.1093/molbev/mst197 PMID: 24132122

33. Hoffmann CS, Winston F. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. Gene. 1987; 57(2): 267–72.

34. Larkin MA, Blackshields G, Brown NP, Chenna R, Mcgettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23(21): 2947–8. PMID: 17846036

35. Altschul S, Madden T, Schaffer A, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids Res. 1997; 25(17): 3389–405. PMID: 9254994

36. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013; 30(12): 2725–9. doi: 10.1093/molbev/mst197 PMID: 24132122

37. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. Nucleic Acids Res. 2008; 36: W197–201. doi: 10.1093/nar/gkn238 PMID: 18463136

38. Martin T, Lu SW, van Tilbeurgh H, Ripoll DR, Dixelius C, Turgeon BG, et al. Tracing the origin of the fungal α1 domain places its ancestor in the HMG-box superfamily: Implication for fungal mating-type evolution. PLoS One. 2010; 5(12).

39. Mallona I, Weiss J, Marcos E-C. pcrEfficiency: a web tool for PCR amplification efficiency prediction. BMC Bioinformatics. 2011; 12(1): 404.

40. Leandro MJ, Sychrová H, Prista C, Loureiro-Dias MC. ZrFsy1, a high-affinity fructose/H+ symporter from fructophilic yeast *Zygosaccharomyces rouxii*. PLoS One. 2013; 8(7): e68165. doi: 10.1371/journal.pone.0068165 PMID: 23279556

41. Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, Van Den Hoff MJ, et al. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Res. 2009; 37(6): e45. doi: 10.1093/nar/gkp045 PMID: 19237396
47. Wolberger C, Vershon AK, Johnson AD, Pabo C. Crystal structure of a MA7α2 homeodomain-operator model complex suggests a general for homeodomain-DNA interactions. Cell. 1991; 67: 517–26. PMID: 1682054

48. Kües U, Casselton LA. Fungal mating type genes—regulators of sexual development. Mycol Res. 1992; 96(12): 993–1006.

49. Anderson JS, Forman MD, Modleski S, Dahlquist FW, Baxter SM. Cooperative ordering in homeo-domain-DNA recognition: solution structure and dynamics of the MA7α1 homeodomain. Biochemistry. 2000; 39(33): 10045–54. PMID: 10955992

50. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009; 25(9): 1189–91. doi: 10.1093/bioinformatics/btp033 PMID: 19151095

51. Belfort M, Roberts RJ. Homing endonucleases: keeping the house in order. Nucleic Acids Res. 1997; 25(17): 3379–88. PMID: 9254693

52. Stoddard BL. Homing endonuclease structure and function. Q Rev Biophys. 2005; 38(1): 49–95. PMID: 16336743

53. Hafiez M, Hausner G, Bonen L. Homing endonucleases: DNA scissors on a mission. Genome. 2012; 55(8): 553–69. doi: 10.1139/g2012-049 PMID: 22891163

54. Gordon JL, Armisen D, Proux-Wera E, O’Eigeartaigh SS, Byrne KP, Wolfe KH. Evolutionary erosion of yeast sex chromosomes by mating-type switching accidents. Proc Natl Acad Sci U S A. 2011; 108(50): 20024–9. doi: 10.1073/pnas.1112808108 PMID: 22123960

55. Ruderfer DM, Pratt SC, Seidel HS, Kruglyak L. 2006 Population genomic analysis of outcrossing and recombination in yeast. Nat Genet. Sep; 38(9):1077–81. Epub 2006 Aug 6 PMID: 16992060

56. Tsai U, Bensonsson D, Burt A, Koutopanou V. Population genomics of the wild yeast Saccharomyces paradoxus: Quantifying the life cycle. Proc Natl Acad Sci U S A. 2008 Mar 25; 105(12):4957–62. doi: 10.1073/pnas.0707314105 Epub 2008 Mar 14. PMID: 18344325

57. Friedrich A, Jung P, Reisser C, Fischer G, Schacherer J. Population genomics reveals chromosome-wide heterogeneous evolution in a protoploid yeast. Mol Biol Evol. 2015 Jan; 32(1):184–92. doi: 10.1093/molbev/msv025 Epub 2014 Oct 27. PMID: 25349286

58. Steensels J, Snoek T, Meersman E, Nicolino MP, Voordeckers K, Verstrepen KJ. Improving industrial production of yeast by functional genomics. Nat Biotechnol. 2014; 32(9): 947–55. doi: 10.1038/nbt.2906

59. Hunter N, Chambers SR, Louis EJ, Borts RH. The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. EMBO J. 1996; 15(7): 1726–33. PMID: 8612597

60. Coyne JA, Orr HA. Speciation. Sunderland, MA. Sinauer Associates, Inc; 2004.

61. Dunn B, Paulish T, Stanbery A, Piotrowski J, Koniges G, Kroll E, et al. Recurrent rearrangement during adaptive evolution in an interspecific yeast hybrid suggests a model for rapid introgression. PLoS Genet. 2013; 9(3): e1003366. doi: 10.1371/journal.pgen.1003366 PMID: 23555283

62. Wolfe KH. Origin of the yeast Whole-Genome Duplication. PLoS Biol. 2015; 13(8): e1002221. doi: 10.1371/journal.pbio.1002221

63. Kinclová O, Potier S, Sychrová H. The Zygosaccharomyces rouxii strain CBS732 contains only one copy of the HOG1 and the SOD2 genes. J Biotechnol. 2001; 88(2): 151–8. PMID: 11403849

64. Baker CR, Booth LN, Sorrells TR, Johnson AD. Protein modularity, cooperative binding, and hybrid regulatory states underlie transcriptional network diversification. Cell. 2012; 151(1): 80–95. doi: 10.1016/j.cell.2012.08.018 PMID: 23021217

65. Wolfe KH. Yesterday’s polyploids and the mystery of diploidization. Nat Rev Genet. 2001; 2(5): 333–41. PMID: 11331899

66. Tuch BB, Galgoczy DJ, Hernady AD, Li H, Johnson AD. The evolution of combinatorial gene regulation in fungi. PLoS Biol. 2008; 6(2): e38. doi: 10.1371/journal.pbio.0060038 PMID: 18303948

67. Dettman JR, Sirjusingh C, Kohn LM, Anderson JB. Incipient speciation by divergent adaptation and antagonistic epistasis in yeast. Nature. 2007; 447(7144): 585–9. PMID: 17124320

68. Brideau NJ, Flores HA, Wang J, Maheshwari S, Wang XU, Barbash DA. Two Dobzhansky-Muller genes interact to cause hybrid lethality in Drosophila. Science. 2006; 314(5803): 1292–5. PMID: 17124320

69. Lee H-Y, Chou J-Y, Cheong L, Chang N-H, Yang S-Y, Leu J-Y. Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species. Cell. 2008; 135(6): 1065–73. doi: 10.1016/j.cell.2008.10.047 PMID: 19070577

70. Bayes JJ, Malik HS. Altered heterochromatin binding by a hybrid sterility protein in Drosophila sibling species. Science. 2009; 326(5959): 1538–41. doi: 10.1126/science.1181756 PMID: 19933102
71. Orr HA, Turelli M. The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. Evolution. 2001; 55(6): 1085–94. PMID: 11475044

72. Piatkowska EM, Naseeb S, Knight D, Delneri D. Chimeric protein complexes in hybrid species generate novel phenotypes. PLoS Genet. 2013; 9(10): e1003836. doi: 10.1371/journal.pgen.1003836 PMID: 24137105

73. Siliciano PG, Tatchell K. Transcription and regulatory signals at the mating-type locus in yeast. Cell. 1984; 37(3): 969–78. PMID: 6378388

74. Harashima S, Miller AM, Tanaka K, Kusumoto K, Mukai Y, Nasmyth K, et al. Mating-type control in Saccharomyces cerevisiae: isolation and characterization of mutants defective in repression by a1-alpha 2. Mol Cell Biol. 1989; 9(10): 4523–30. PMID: 2685555

75. Strathern J, Shafer B, Hicks J, McGill C. a/α-specific repression by MATα2. Genetics. 1988; 120(1): 75–81. PMID: 2851482

76. Goossens K, Ielasi F, Nookaew I, Stals I, Alonso-Sarduy L, Daenen L, et al. Molecular mechanism of flocculation self-recognition in yeast and its role in mating and survival. mBio 6(2): e00427–15. doi: 10.1128/mBio.00427-15 PMID: 25873380

77. Sternberg PW, Stern MJ, Clark I, Herskowitz I. Activation of the yeast HO gene by release from multiple negative controls. Cell. 1987; 48(4): 567–77. PMID: 3545494

78. Fabre E, Muller H, Therizols P, Lafontaine I, Dujon B, Fairhead C. Comparative genomics in hemiascomycete yeasts: evolution of sex, silencing, and subtelomeres. Mol Biol Evol. 2005; 22(4): 856–73. PMID: 15616141

79. Rajaei N, Chiruvella KK, Lin F, Aström SU. Domesticated transposase Kat1 and its fossil imprints induce sexual differentiation in yeast. Proc Natl Acad Sci U S A. 2014; 111(43): 15491–6. doi: 10.1073/pnas.1406027111 PMID: 25313032

80. Mori H, Onishi H. Diploid hybridization in a heterothallic haploid yeast, Saccharomyces rouxii. Appl Microbiol. 1967; 15(4): 928–34. PMID: 6049309

81. Marcet-Houben M, Gabaldón T. Beyond the Whole-Genome Duplication: phylogenetic evidence for an ancient interspecies hybridization in the baker’s yeast lineage. PLoS Biol. 2015; 13(8): e1002220. doi: 10.1371/journal.pbio.1002220 PMID: 26252497