



**HAL**  
open science

# Multiple parameters drive the efficiency of CRISPR/Cas9-induced gene modifications in *Yarrowia lipolytica*

Vinciane Borsenberger, Djamila Onesime, Delphine Lestrade, Coraline Rigouin, Cécile Neuvéglise, Fayza Daboussi, Florence Bordes

## ► To cite this version:

Vinciane Borsenberger, Djamila Onesime, Delphine Lestrade, Coraline Rigouin, Cécile Neuvéglise, et al.. Multiple parameters drive the efficiency of CRISPR/Cas9-induced gene modifications in *Yarrowia lipolytica*. *Journal of Molecular Biology*, 2018, 430 (21), pp.4293-4306. 10.1016/j.jmb.2018.08.024 . hal-02154346

**HAL Id: hal-02154346**

**<https://hal.science/hal-02154346>**

Submitted on 26 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



# Multiple Parameters Drive the Efficiency of CRISPR/Cas9-Induced Gene Modifications in *Yarrowia lipolytica*

Vinciane Borsenberger<sup>1</sup>, Djamila Onésime<sup>2</sup>, Delphine Lestrade<sup>3</sup>, Coraline Rigouin<sup>1</sup>, Cécile Neuvéglise<sup>2,†</sup>, Fayza Daboussi<sup>1,†</sup> and Florence Bordes<sup>1,†</sup>

**1** - LISBP, Université de Toulouse, INSA, INRA, CNRS, Toulouse, France

**2** - Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, Paris, France

**3** - TWB, Université de Toulouse, CNRS, INRA, INSA, Toulouse, France

**Correspondence to Florence Bordes:** LISBP, Université de Toulouse, INSA, INRA, CNRS, 135 avenue de Rangueil, 31077, Toulouse CEDEX 04, France. [bordes@insa-toulouse.fr](mailto:bordes@insa-toulouse.fr)

<https://doi.org/10.1016/j.jmb.2018.08.024>

Edited by Dylan Taatjes

## Abstract

*Yarrowia lipolytica* is an oleaginous yeast of growing industrial interest for biotechnological applications. In the last few years, genome edition has become an easier and more accessible prospect with the world wide spread development of CRISPR/Cas9 technology. In this study, we focused our attention on the production of the two key elements of the CRISPR–Cas9 ribonucleic acid protein complex in this non-conventional yeast. The efficiency of NHEJ-induced knockout was measured by time-course monitoring using multiple parameters flow cytometry, as well as phenotypic and genotypic observations, and linked to nuclease production levels showing that its strong overexpression is unnecessary. Thus, the limiting factor for the generation of a functional ribonucleic acid protein complex clearly resides in guide expression, which was probed by testing different linker lengths between the transfer RNA promoter and the sgRNA. The results highlight a clear deleterious effect of mismatching bases at the 5' end of the target sequence. For the first time in yeast, an investigation of its maturation from the primary transcript was undertaken by sequencing multiple sgRNAs extracted from the host. These data provide insights into of the yeast small RNA processing, from synthesis to maturation, and suggests a pathway for their degradation in *Y. lipolytica*. Subsequently, a whole-genome sequencing of a modified strain detected no abnormal modification due to off-target effects, confirming CRISPR/Cas9 as a safe strategy for editing *Y. lipolytica* genome. Finally, the optimized system was used to promote *in vivo* directed mutagenesis *via* homology-directed repair with a ssDNA oligonucleotide.

© 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## Introduction

With interesting characteristics such as its high capability at lipid accumulation [1] and protein secretion [2], *Yarrowia lipolytica* is an attractive organism to study and engineer in the view of synthetic biology applications [2–6]. It is categorized as a “generally recognized as safe” microorganism [7] and its genome is entirely sequenced, which makes it an ideal candidate for genome editing studies. Recently, the CRISPR/Cas9 system from *Streptococcus pyogenes* [8–10] has successfully been used to edit its genome [11–15]. The engineered CRISPR/Cas9 system uses a nuclease (Cas9) acting in complex with a single-guide RNA

(sgRNA) to generate a double-strand break (DSB) at a targeted locus [9]. The specificity of the ribonucleic acid protein (RNP) complex for its target stems from the recognition by the sgRNA of a ~20-b sequence complementary of a spacer sequence at the 5' end of the guide through Watson Crick base pairing. The process requires the target sequence in the genome to be next to a protospacer adjacent sequence (PAM) to enable Cas9 fixation—with the PAM being NGG in *S. pyogenes* system. The DSB is then repaired by the non-homologous end-joining (NHEJ) mechanism, which can introduce small indels (insertions/deletions) at the repair site, which results in reading frameshifts [11] leading to gene inactivation. Furthermore, when the

CRISPR system was used in conjunction with a DNA matrix bearing homologous arms to the DSB region, the homologous recombination at the break site was demonstrated to be drastically enhanced in *Y. lipolytica* and applied for easy gene integration strategies [15].

Currently, the most commonly used methodology for genome edition in *Y. lipolytica* relies on an expression system adapted to this particular yeast. The nuclease expression is usually driven by a very strong constitutive promoter such as the hybrid 8UAS-pTEF, while the sgRNA is produced under the control of a transfer RNA (tRNA) glycine [11]. The use of a tRNA glycine as a promoter for sgRNA was originally based on the expression system of 5S RNA units [16], since in *Y. lipolytica* genome multiple copies of this small functional RNA can often be found 4 to 5 bases directly downstream of a tRNA gene. The system recruits the polymerase III (Pol III), specialized in the production of small non-coding RNA by means of conserved sequences within the tRNAs, and stops the transcription at stretches of Ts. The 5S RNA units are produced in fusion with the tRNA glycine in a single transcript, to be later processed by the RNase Z [17]. However, little is known about the precise maturation process and the resulting end products in *Y. lipolytica*. In this paper, we address this issue by sequencing the 3' and 5' ends of the sgRNA produced *in vivo* by the yeast.

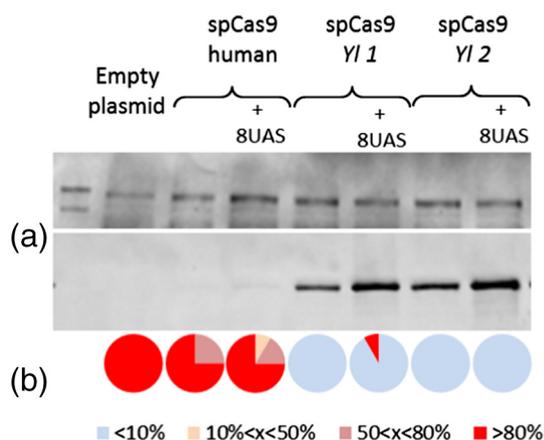
Consequently, as CRISPR/Cas9 presents so much promise for the genome edition of *Y. lipolytica*, it appeared of crucial interest to conduct a thorough investigation on the impact of factors such as nuclease expression levels and sgRNA maturation on DSB efficiency. To this effect, the quantitative evaluation of CRISPR/Cas9 system in *Y. lipolytica* was performed using the Y6047 strain, which has a single copy of the RedStar2 gene under the TEF promoter (without its intron) integrated stably in one copy in the genome. RedStar2 is a DsRed mutant whose sequence harbors several mutations enabling its rapid maturation [18] and was optimized for expression in *Y. lipolytica* [19], hence allowing for a fast quantification of complete or partial gene disruption by fluorescence monitoring. The data we obtained give a clear assessment of the impact of Cas9 production levels in the yeast and of the influence of the sgRNA directly upstream sequences on NHEJ-induced knockout (KO). Moreover, it provides a better understanding of the functioning of Pol III driven sgRNA expression in *Y. lipolytica*. The optimized system achieves high levels of effectiveness, with whichever of the three most commonly used selection markers in this yeast, and is applied to editing the endogenous gene *FAD2* by homology-directed repair (HDR) with ssDNA oligonucleotide.

## Results

In order to analyze Cas9 expression in *Y. lipolytica*, single plasmids were generated with different codon

optimizations for Cas9 expression (see SI Protocols). The first was a human codon bias sequence that had been successfully used to engineer *Saccharomyces cerevisiae* genome [20]. The second was an in-house design using *Y. lipolytica* codon bias (termed spCas9 YI1; see SI Fig. S1), which we compared to a third published optimized sequence [11] (termed spCas9 YI2) that shares a 87% nucleotide identity with spCas9 YI1. The three gene versions were placed under the control of two constitutive promoters: the natural pTEF (strong) and the hybrid 8UAS-pTEF [21] (very strong). None of these included the TEF intron. All sequences harbored the SV40 nuclear localization signal at their C terminals. The six constructs were cloned in the same backbone plasmid comprising of the sequence for sgRNA expression (see SI Fig. S2), the autonomously replicating sequence ARS 68 [22], the resistance gene to hygromycin hph [23] from *Escherichia coli*, and a shuttle part for production in *E. coli*, originally found in pBR322 (see SI Fig. S3A). The Y6047 strain was transformed with each of the six plasmids using the antibiotic selection and plated out. Colonies were then cultivated in YPD supplemented with hygromycin. Toward the end of the log phase, the cells were collected and lysed, and the cellular extracts were analyzed by Western blot with Cas9-specific antibodies (Fig. 1(a)). In parallel, 12 individual colonies per transformation were cultivated in rich medium (YPD), and the ratio of fluorescence over OD was taken when the stationary phase had been reached (Fig. 1(b)). The Western blot demonstrates that Cas9 is produced with every plasmid, albeit at very low levels when using a human codon bias, while both *Y. lipolytica* optimized sequences produced over 50-fold more chemiluminescence signals. In addition, the use of the 8UAS sequences enhances the levels of detected Cas9 by 2- to 3-fold with every gene version. Similar results are observed when using YNB medium for growth (see SI Fig. S5). Consistently with the detected nuclease levels, no culture from the colonies transformed with the human codon bias Cas9 showed a complete inactivation of the fluorescence. Although 25% of them produced a diminished fluorescence signal indicative of the presence of disrupted strains within the population (see SI Fig. S4), overall, most of the colonies behaved like the parent strain. In contrast, colonies from the transformations with the *Y. lipolytica* optimized nucleases produced over 90% RedStar2 disrupted strains with no measurable fluorescence levels over background, with or without 8UAS enhancement.

To analyze the kinetic of targeted mutagenesis and impact of Cas9 expression on cell viability, we examined the early generation of RedStar2 disrupted strains, in cells transformed by CRISPR/Cas9 plasmids using multiparameter flow cytometry. Plasmids were generated with the *Y. lipolytica* optimized spCa9 YI1, with or without 8UAS, with the sgRNA complete with or without the RedStar2 target (see SI Protocols). The four constructs were cloned in the same backbone

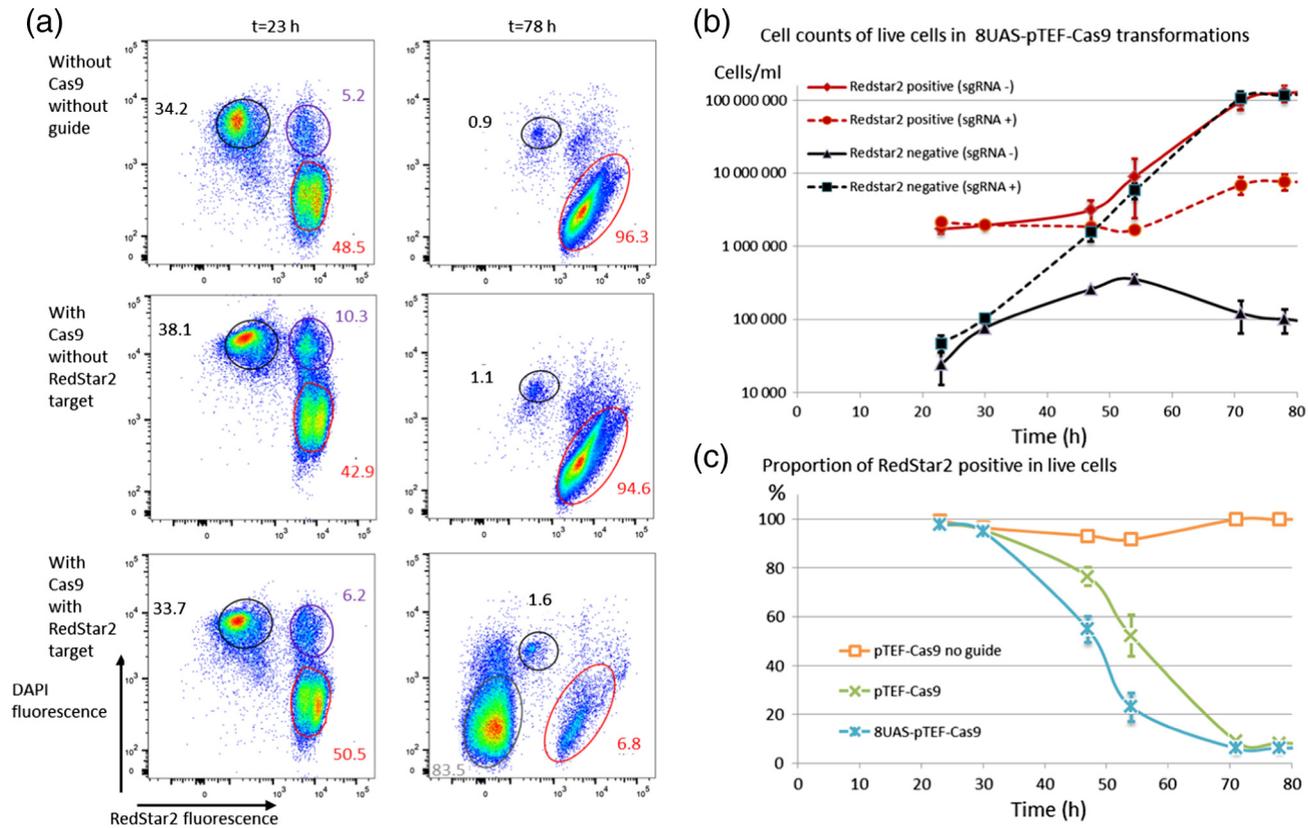


**Fig. 1.** Influence of Cas9 expression levels on RedStar2 disruption efficiency. Line 1: size marker, line 2: empty plasmid, lines 3 and 4: Cas9 with a human codon bias, lines 5 and 6: Cas9 with *Y. lipolytica* codon bias *YL1*, lines 7 and 8: comparison with *Y. lipolytica* codon bias *YL2*. In lines 3, 5, and 7, a strong constitutive promoter is used (pTEF); in lines 4, 6, and 8, a stronger hybrid promoter is used (8UAS–pTEF). (a) Western blot of YPD cultivated *Y. lipolytica* colonies of the Y6047 strain transformed with plasmids expressing Cas9. The membrane was visualized first using the Stain Free technology (top panel) to control the membrane protein loading, then with monoclonal anti-Cas9 mouse specific antibodies (bottom panel). (b) Fluorescence repartition in Y6047 strain colonies exposed to CRISPR/Cas9 with a sgRNA targeting RedStar2, measured on 12 individual colonies per transformation: ratio of fluorescence over OD<sub>600</sub> was quantified and compared to the ratios observed in the original strain. Blue indicates that less than 10% of the original fluorescence is retained; light pink, between 10% and 40%; dark pink, between 50% and 80%; and red, over 80%.

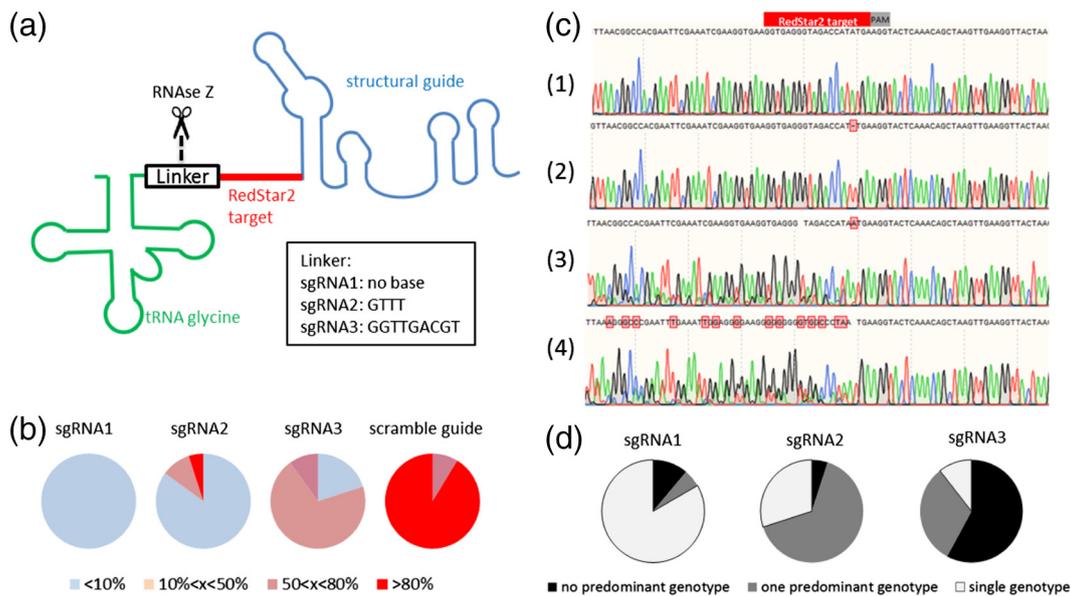
plasmids comprising, the autonomously replicating sequence ARS 18 [22], a LEU2 auxotrophic marker, and a shuttle part for production in *E. coli* (see SI Fig. S3B). The purpose of swapping to a smaller replication origin was to assess the possibility of using more compact plasmids. Although this change did not seem to have a measurable impact on transformation efficiency, using the LEU2 selection marker raised it by more than 100-fold (data not shown). Y6047-competent cells were transformed with the plasmids, as well as control plasmids without Cas9 or sgRNA sequence, and grown directly, without plating, in YNB liquid medium supplemented with glucose and uracil. Aliquots were taken at intervals for 4 days and treated with DAPI prior to being injected in the flow cytometer. DAPI enters senescent and damaged cells faster than healthy cells, allowing for the differentiation of fit CRISPR transformants from dying plasmid-free cells (Figs. 2(a) and S6). Monitoring of the RedStar2 fluorescence of the healthy cell population enables the direct visualization of gene disruption inside the CRISPR active transformations over time. In the early stages, all transformations display similar behavior, with two main populations, one made of DAPI-negative

and Redstar2-positive cells (red circle–“healthy cells”) and the other being made of DAPI sensitive and Redstar2 negative cells (black circle–senescent cells). All the transformations show a lag phase spanning from 30 to 47 h (see SI Fig. S6), during which the population of DAPI-positive cells represents 40% to 50% of the considered events. In contrast, all DAPI-negative cells exhibit strong Redstar2 fluorescence. After 78 h of culture, “healthy” cells with a low DAPI uptake have taken over, eventually accounting for more than 90% of the population. In control transformations lacking a targeting guide to RedStar2 or Cas9, all these cells also display RedStar2 fluorescence, while in transformations using Cas9 with a targeting guide, RedStar2-negative cells are predominant (gray circle). In the latter, cell count monitoring over time shows that the cell growth of RedStar2 negative is analogous to the cell growth of Redstar2 positive in the controls and clearly indicates that disrupted cell lines are generated within 30 h of growth (Fig. 2(b)) and demonstrates the lack of toxicity induced by Cas9 expression. Conversely, no progression of the Redstar2 positive is perceived for 50 h, after which a small growth can be observed. Finally, transformations with a pTEF–Cas9 and 8UAS–pTEF–Cas9 produced RedStar2 disrupted cells in comparatively similar amount of time (Fig. 2(c)), and in every culture after 70 h, the proportion of disrupted over non-disrupted cells showed little changes.

Then, we focused our study on the expression of the sgRNA driven by RNA Pol III, which was based originally on the expression system of 5S RNA units. Therefore, we aligned the sequences of the 28 tRNAglycine–5SRNA associations that can be found in the genome [16] and obtained a strong consensus sequence of 4 bases (see SI Fig. S7). We compared the efficiency of three tRNAgly–sgRNA association with different linkers to bind them together. The three constructs had either no bases, the 4 bases from the consensus sequence, or a 9 bases sequence used in previous report [11] (Fig. 3(a)). The three tRNAgly–sgRNA sequences were placed in a pTEF–Cas9 plasmid, to be transformed in Y6047 strain using hygromycine as a selection marker. Individual colonies picked from the plated transformations were then cultivated in YPD without antibiotic. Ratio of fluorescence over OD was taken when the stationary phase had been reached to be compared to that of the parent strain (Fig. 3(b)). In parallel, aliquots of the cultures were pelleted and lysed in alkaline buffer. Lysis extracts were analyzed by PCR at the targeted locus and sequencing (Fig. 3(c) and SI Table 1). The sequence traces were analyzed with the TIDE software to identify and quantify the frequency of genotypes within the population [24]. Comparison between 35 control colonies at the RedStar2 locus displayed 95% to 98% correspondence to the wild-type (WT) sequence. Every colony transformed with the direct tRNA–sgRNA association with no intervening base had completely



**Fig. 2.** Multiparameter FLOW CYTOMETRY monitoring of transformation of the Y6047 strain with CRISPR/Cas9 plasmids targeting RedStar2. (a) Cells repartition as a function of their DAPI and RedStar2 fluorescence after transformation with an empty vector (upper panels) or pTEF-Cas9 vectors without targeting (middle panels) or with a sgRNA targeting RedStar2 (bottom panels). Sub-type populations are identified with circles (black: DAPI-positive/RedStar2 negative; purple: DAPI-positive/RedStar2 positive; red: DAPI negative/RedStar2 positive; gray: DAPI negative/RedStar2 negative). The percentage of each group relative to the totality of considered events is indicated next to their respective circle. (b) Counts of RedStar2-positive and -negative cells depend on the presence of a targeting guide in the transformation with 8UAS-pTEF-Cas9 vectors. (c) Percentage of RedStar2-positive cells within the DAPI-negative population. Each vector was tested in three biological replicates. (b and c) Error bars reflect the standard deviation from the mean of these replicates.



**Fig. 3.** 5' Sequence influence on RNA guide efficiency. (a) Schematic representation of sgRNA excision from a primary transcript by RNase Z and sequences tested as linkers between the tRNA and the sgRNA sequences. (b) Fluorescence repartition in the Y6047 strain colonies exposed to CRISPR/Cas9 with a sgRNA targeting RedStar2 expressed with different linker lengths: ratio of fluorescence over  $OD_{600}$  was measured and compared to the ratios observed in the original strain. The experiment was performed on 20 individual colonies for each plasmid. Blue indicates that less than 10% of the original fluorescence is retained; light pink, between 10% and 40%; dark pink, between 50% and 80%; and red, over 80%. (c) Typical chromatograms from sequenced colonies aligned with WT sequence (1), (2) unique genotype mostly found in colonies transformed with sgRNA1. TIDE indicates that the genotype with a 1-base deletion represents 97% of this trace decomposition (3): mixed genotypes with a predominant one (in this instance, the TIDE score is 66.8% for the A insertion genotype), (4) mixed genotype, typical in colonies transformed with sgRNA3: here, the most frequent genotype represents 28.5% of the trace. (d) Unicity of recovered genotypes as a function of the sgRNA used. Light gray refers to colonies where the predominant genotype represents more than 95% of the trace decomposition; dark gray, colonies with a predominant genotype representing between 50% and 95% of the trace decomposition; and black, colonies where the most frequent genotype does not exceed 50%.

lost its fluorescence, and genomic analysis showed that most of them were composed of one KO genotypes representing more than 95% of the trace decomposition according to TIDE (Fig. 3(d)), which we assimilated as “unique” genotype. In contrast, colonies transformed with the 9-base linker presented a pronounced mosaicism, visible at the phenotypic level, with most of the colonies displaying a diminished fluorescence signal-marking them as a mixture of both WT and disrupted cell lines. Accordingly, the intermediate association that used a 4-base linker produced a result in between, with the mosaicism being only evident at the genotypic level, with entirely non-fluorescent colonies composed of disrupted cell lines with different indels. To further investigate these results, in a separate experiment, we compared the fluorescence levels of cultures while maintaining the selection pressure on the CRISPR system with the fluorescence of the same colonies grown in the absence of hygromycin and found little differences in the phenotypic results (see SI Fig. S8). In addition, since the Scr1 sequence was shown to enhance CRISPR/Cas9 activity when placed upstream of the tRNA<sup>gly</sup>–sgRNA association [11], this improved promoter version was tested with the three linkers (see SI Fig. S3C). The

corresponding plasmids were transformed in strain Y6047 and the fluorescence measurements, as well as sequencing analysis, demonstrated an improvement of the RNP complex efficiency with all three linkers. Colonies transformed with the direct Scr1–tRNA–sgRNA association were more often made of a single genotype, and colonies transformed with the corresponding sequence with the 4- and 9-base linkers were more often fully disrupted for RedStar2. In addition, the combined sequencing results show that in single genotype colonies and colonies with a dominating genotype, the majority of repair events are simple indels (see SI Table 1). The most frequent occurrence is the deletion of one base (52%) or insertion of one base (32%, almost exclusively A), occurring 3 bases upstream of the PAM, which is typical of mutations induced by CRISPR/Cas9 through NHEJ repair. Of note, a few long insertions were found and could be traced back to *Salmo salar*. We can assume that these sequences were acquired from salmon sperm present in the transformation as DNA carrier. Taken together, all these results demonstrate that having no linker upstream of the sgRNA sequence produces more efficient and rapidly acting RNP complex.

The direct tRNA–sgRNA association with no intervening base proved the best suitable candidate for efficient CRISPR/Cas9 editing; therefore, we proceeded to the 5′ and 3′ ends mapping of sgRNA produced using this configuration. To this effect, cells were transformed with the plasmids from the linker impact on the RedStar2 disruption study harboring the direct tRNA–sgRNA association (see SI Fig. S3C), harvested in early log phase, and then their RNA was isolated. The 5′ and 3′ ends of sgRNAs were sequenced using PCR on circularized RNA (cRT-PCR) (Fig. 4A) [17,25]. In this procedure, RNA extracts are circularized prior to being reverse transcribed to cDNA using a sgRNA specific primer. Following this, PCR using primers to amplify the section across the circularization junction gave access to both 3′ and 5′ ends of the sgRNA. We found a variety of sequences reflecting the life cycle of sgRNAs in *Y. lipolytica* (Fig. 4B and see SI Table 2). Some were found still attached to tRNAgly, not yet processed by RNase Z. Others were neatly cut at the 3′ end of the tRNA, giving the expected full-length target sequence, and others lacked a few bases—the most frequent occurrence in incomplete guides being spacers of 15 bases. Interestingly, a number of these sgRNA had polyA tails appended at their 3′ end, indicative of a possible degradation pathway [26–30].

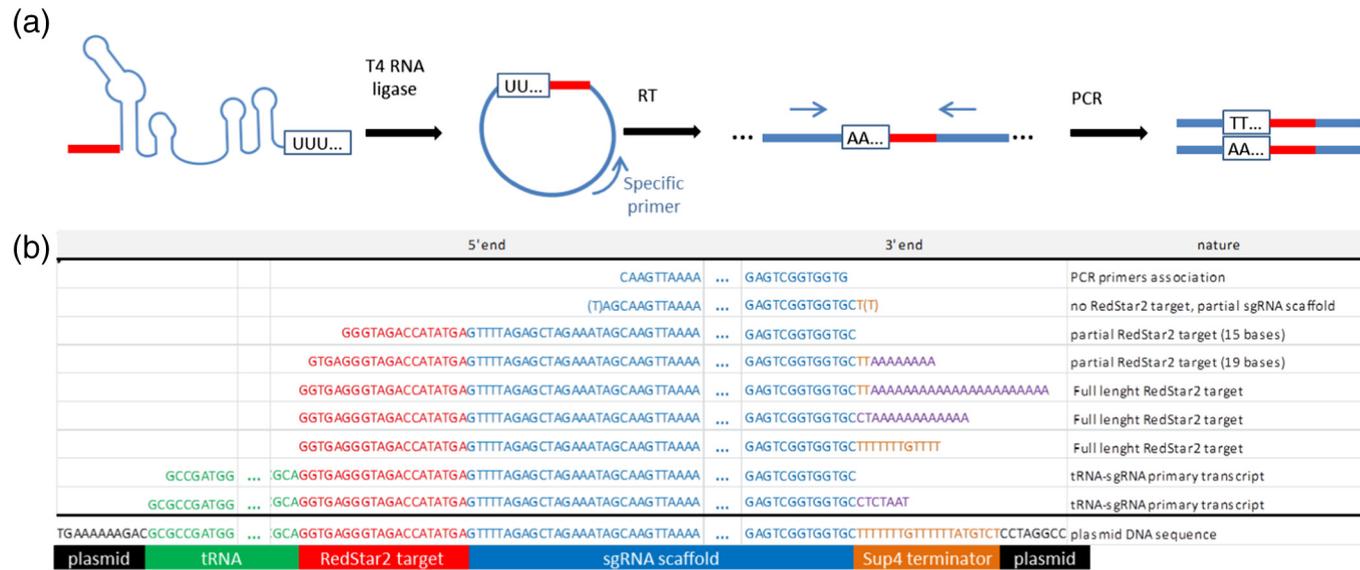
Finally, in order to evaluate the frequency of off-target activity in *Y. lipolytica*, we sequenced and compared the genome of strain Y6047 and six CRISPR-modified strains. Strain Y6047 was transformed with three plasmids containing either no sgRNA (pCas9-URA3), a tRNAgly–sgRNA (pCg53-DsRed), or a ScR1–tRNAgly–sgRNA (pCg58-DsRed) (see SI Fig. S3D). The three constructs were cloned in the same backbone plasmids comprising the autonomously replicating sequence ARS 18, a URA3 auxotrophic marker, and a shuttle part for production in *E. coli*. Two single colonies from each transformation were selected for subsequent plasmid loss and genome sequencing. Y6047 genome was assembled into 27 scaffolds larger than 1 kb (see SI Table S4), and no major chromosomal rearrangements were observed compared to the WT parent strain W29 [31] (see SI Fig. S9A). Similarly, a strong collinearity was observed between the genomes of Y6047 and CRISPR-modified strains. In strain Y6047, the RedStar2 gene was located in scaffold24, in the intergenic region between YAL10F19998g and YAL10F20020g (see SI Fig. S9B). As expected, no mutation in RedStar2 was observed for both isolates transformed with pCas9-URA3, whereas the four other transformants transformed with the equivalent plasmids with a sgRNA showed indel events that occurred 3 bases upstream of the PAM: insertion of an A in Y6047-p53\_1, deletion of an A in Y6047-p53\_2 and Y6047-p58\_2, and deletion of AT in Y6047-p53\_1. Sequencing reads were then mapped to the Y6047 assembled genome to detect SNPs and

indels, which were subsequently filtered for depth and frequency. It is frequent that sequence errors occur during the assembly steps especially in repeated regions, as seen in the alignment of Y6047 reads with its genome assembly (see SI Table 5). Interestingly, transformants had a raw number of SNPs and indels in the same range as Y6047, and filtered positions were mainly the same as in Y6047. Finally, the number of unique mutations (SNPs plus indels) in transformants was around 10, with most of them occurring in microsatellites or palindromic sequences, suggesting sequencing errors rather than true mutations. Knowing that the genome size of *Y. lipolytica* is 20.5 Mb, we thus can consider the off-target activity as insignificant in this yeast.

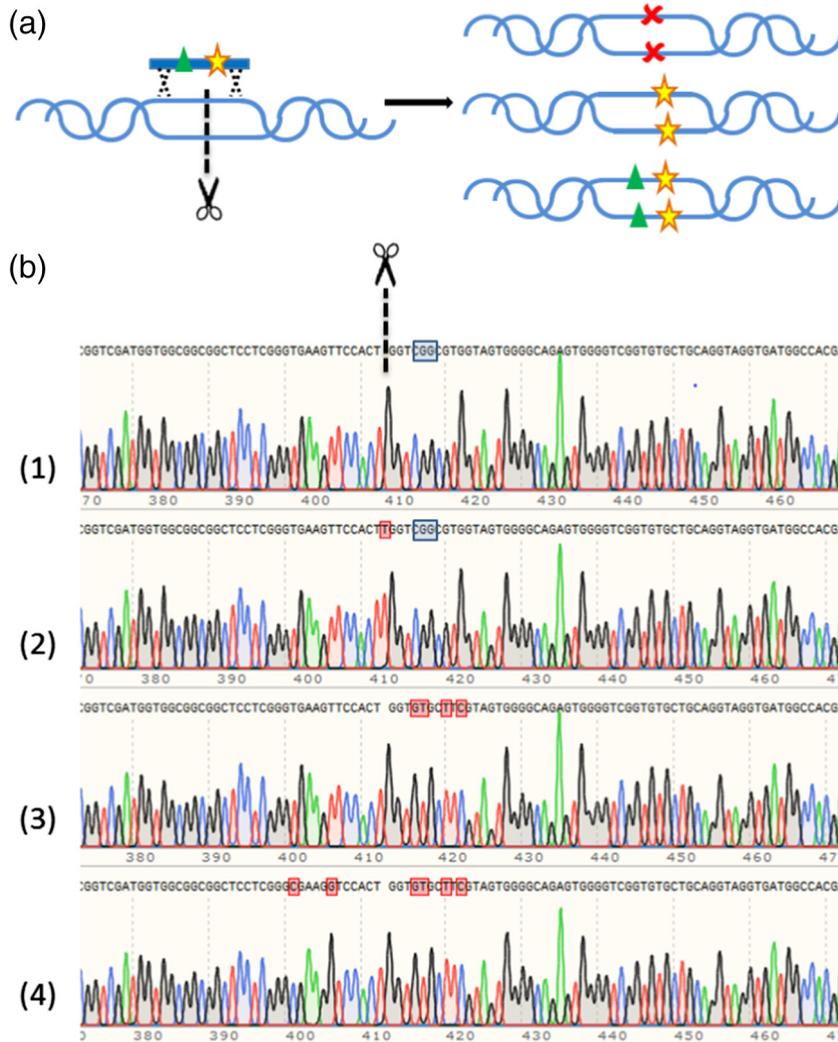
Having established the lack of off-target effects of CRISPR/Cas9 use in *Y. lipolytica*, we tested its ability to promote homology directed repair (HDR) with a ssDNA oligonucleotide as repair template [32,33]. To this effect, a target sequence in a conserved region of the *FAD2* gene [34] (YAL10B10153g) was designed and integrated in the generic plasmid pCg58 (see SI Protocols), with pTEF-spCa9 Y11, ScR1-tRNAgly, and the URA3 marker. The plasmid was transformed into *Y. lipolytica* strain Po1d [35] alongside a 90b-ssDNA oligonucleotide homologous to the target region with six mutations from the WT sequence straddling the expected DSB location (see Fig. 5(a) and SI Table 3). Both 3′ and 5′ ends of the oligonucleotide were protected by phosphothioate linkages in order to improve their lifetime in the host [36] (see SI Table 3). Transformants were isolated after 3 days and analyzed by PCR and sequencing (see Table 1). Trace decomposition was performed using TIDER [37] in order to quantify the introduction of the desired point mutations. In 66% of the colonies, the estimates showed single genotypes with frequencies above 97%. Moreover, all of the colonies transformed with the CRISPR/Cas9 plasmid with the sgRNA had a dominant genotype with a modified genome (i.e., above 50% frequency in the trace, according to TIDER; see Table 1). As seen previously during the RedStar2 disruption study, small deletions were found to be the most frequent event (9 out of 11 colonies when the plasmid was transformed alone, and 18 out of 36 when the oligonucleotide was also present). Two colonies had longer insertions: a 93-bp fragment of salmon DNA and a 46-bp fragment of unknown origin. Remarkably, in transformations with the oligonucleotide, we found one occurrence in which the expected mutations appeared on only one side of the DSB (see Fig. 5), and, even more remarkably, 6 out of 36 of the colonies harbored the desired mutations.

## Discussion

The efficiency of the CRISPR/Cas9 system in any host is dependent on the quantity of adequately formed



**Fig. 4.** Sequences of sgRNA ends. (a) Principle of cRT-PCR: small RNAs are circularized first, then reverse transcribed. PCR is performed on the cDNA using primers that amplify the 5' and 3' ends of the sgRNA, which are now facing each other. (b) Examples of sequences obtained from TOPO cloning of the cRT-PCR of sgRNA. The bases were color coded to help visualization: green, tRNA sequence; red, RedStar2 target; blue, sgRNA scaffold; orange, Pol III terminator; purple, bases added *in vivo*.



**Fig. 5.** Directed mutagenesis using ssDNA oligonucleotide (a) Principle of CRISPR/Cas9 promoted HDR: a DSB is generated, and concomitantly a repair template with 40-base homology sequences on both sides of the break site is injected. Possible outcomes for the host include, for instance, the introduction of indels by the NHEJ repair mechanism (top) and the introduction of the mutations from the repair template on only one side of the DSB (middle) or both sides (bottom). (b) Trace (1) is a sequence chromatogram obtained from colonies transformed with the ssDNA oligonucleotide and Cas9 without the targeting sgRNA. Traces (2), (3), and (4) are obtained from colonies transformed with the ssDNA oligonucleotide and Cas9 with the targeting sgRNA. The PAM is squared in a blue box, and mutations are indicated in a red box. Trace (1) shows a WT genotype; trace (2) shows the introduction of a T at the break site by NHEJ; trace (3) illustrates the insertion of mutations from the ssDNA oligonucleotide on one side of the PAM and trace (4) on both sides.

RNP complex, a parameter that is both influenced by Cas9 and sgRNA expressions. In *Y. lipolytica*, when using a human codon bias sequence, we saw that while Cas9 is produced at very low levels, some gene disruptions are still detectable. In addition, overexpression of an optimized codon bias Cas9 using the TEF promoter was enough to obtain a full disruption of the RedStar2 gene and the 8UAS enhancement appeared to be unnecessary. This is further illustrated in the multiparameter flow cytometry data where we obtained identical proportions of RedStar2 negative over positive in both pTEF-Cas9 and 8UAS-pTEFCas9 transformations at the end of the growth phase. By contrast with the swift occurrence of disrupting mutations events in the early log phase, the RedStar2-positive counts in CRISPR active transformation only rose after a 50-h lag phase. This is in accordance with previous observations in *Schizosaccharomyces pombe*, showing that the survival of the transformants with intact WT genome

was impaired by the presence of an active CRISPR system, ultimately preventing cell division and leading to either cell death or CRISPR system inactivation [38]. Hence, RedStar2-positive clones can only stem from replication of WT cells at very slow rates, cell lines with disrupting mutations in CRISPR/Cas9 system, or cell lines with modified genotypes without out-of-frame mutations. In *Y. lipolytica* this toxicity toward WT genome clones was not observed in either pTEF-Cas9 or 8UAS-pTEF-Cas9 transformations, as the growth in CRISPR/Cas9-modified cells was essentially identical to that of the ones observed in control transformations. Nevertheless, Cas9 overexpression with the strong promoter  $P_{TDH3}$  was reported to induce toxicity in *S. cerevisiae*, which was circumvented by the use of the weaker  $P_{RNR2}$  promoter [39]. This information along with the observed toxicity of spCas9 in *S. pombe* [38] lead us to favoring the TEF promoter over the stronger hybrid 8UAS-pTEF, since the former

**Table 1.** Dominant genotypes in individual colonies sequenced at the *FAD2* locus after transformation with the CRISPR/Cas9 plasmids with or without ssDNA-oligonucleotide and with or without a targeting sgRNA

	Cas9 + sgRNA	Cas9 + sgRNA + ssDNA oligonucleotide	Cas9 + ssDNA oligonucleotide
Deletion –1 base	3	14	
Deletion –2 bases		1	
Deletion –3 bases	2		
Deletion –5 bases	2	2	
Deletion –8 bases	1		
Deletion –19 bases	1	1	
Insertion T	2	7	
Insertion G		1	
Exogenous DNA insert		2	
HDR on one side of the DSB		1	
HDR on both sides of the DSB		6	
Other indel event		1	
WT			4
Total colony screened	11	36	4

In all colonies, a dominant genotype was found to represent more than 50% of the trace according to TIDER.

has the advantage of being less taxing on the yeast resources and reduces the potential harmfulness of the nuclease, without significantly weakening its efficiency.

Hence, with our existing set of *Y. lipolytica* genetic tools, the current limiting factor for the generation of functional RNP complex clearly resides in guide expression. We observed that using a linker between the tRNA promoter and the sgRNA drastically diminishes the efficiency of the RNP complex, which can most likely be attributed to the presence of additional bases in the 5' end of the target sequence in the sgRNA causing mismatches with the genomic sequence, hence reducing the affinity for the target. In contrast, when the guide is properly designed, gene KO happens so fast that a single genotype per colony is detected (Figs. 2 and 5). Likewise, single-genotype colonies of mutated strains are obtained when the optimized system is employed in the editing of *FAD2*, an oleate delta-12 desaturase, in which disruption results in the obliteration of linoleic acid production in *Y. lipolytica* [34]. Similar proteins producing different fatty acid profiles with varying degree of identity exist in nature [40,41]. A thorough assessment of the impact of residue variations in conserved regions in relation to fatty acid production presents a great interest at fundamental and applied levels. Therefore, the possibility of performing directed mutagenesis *in vivo* in order to be able to quickly generate multiple mutants of this protein is attractive and the optimized CRISPR/Cas9 system provides a simple, fast, and easy way to achieve it. Although a high homologous recombination ratio has been achieved using large recombination matrix of 1 kb or more [12,15], the ssDNA oligonucleotide HDR, if less effective, bypasses the homology matrix assembly step.

In addition, cRT-PCR showed the presence of shorter guides, with trimmed bases in the target

sequence. This finding may explain that the 4-base linker still enables the CRISPR/Cas9 system to be functional, even if less efficient. A question still remains, as whether they originated from the degradation of mature full-length sgRNA—occurring *in vivo* or during the extraction process, or were cleaved from the primary transcript in this fashion. The latter would imply that the maturation process is not very specific in *Y. lipolytica*, unlike what is observed in rice [17]. At any rates, some of the shorter sgRNA would probably still be functional, as Fu and co-workers [42] demonstrated that guides with target sequences of 17 bases were generally as effective as 20-base ones, with better selectivity. The same study also found that shorter RNA guides with sequences of 16 bases and below were ineffective. Therefore, in theory, the number of potential off-target sites that would result from the presence of shorter guides is limited to sequences with a 17-base homology to the target next to a PAM sequence in the genome. As the region 17 to 20 bases from the PAM is already well known for tolerating mismatches [9], such sequences would already have been flagged as potential off target sites. In addition, we confirmed the specificity of the CRISPR/Cas9 system by the complete genome sequencing of CRISPR-modified Y6047 strain, which did not reveal major perturbations resulting from the transient expression of CRISPR/Cas9, deeming the whole system safe to use.

Interestingly, a number of sgRNA was found with polyA tails showing that some had been polyadenylated in the absence of a polyA signal. It is worth noting that in other hemiascomycetes, the TRAMP complex [26–30] in charge of RNA quality control and viral and parasitic RNA repression initiates the aberrant RNA removal process by a polyadenylation step, directing them to the nuclear exosome, where they are subsequently degraded. In *Y. lipolytica*, this step could be undertaken by the protein encoded by

YALI0F30525g, which shows 46% similarity over 642 amino acids with the TRF5 of *S. cerevisiae* [43], a non-canonical poly(A) polymerase that polyadenylates misfolded tRNAs, snoRNAs, and pre-rRNAs, as part of the TRAMP complex. Hence, the polyA tail suggests that sgRNAs are efficiently eliminated by the host. In this regard, we understand that being able to strongly express sgRNAs in order to balance their degradation is a key factor for an efficient CRISPR/Cas9 editing.

## Conclusion

We hope that we provided some clear insights into the use of CRISPR/Cas9 system in *Y. lipolytica* with a systematic approach, considering both aspects of the RNP formation. Our study clearly establishes that CRISPR/Cas9-induced KO is a very fast-acting process and demonstrates that one should preferentially opt for the direct attachment of the sgRNA to the tRNA glycine sequence to benefit from the best efficiency. We put the optimized system to the test by delivering the first example of *in vivo* directed mutagenesis in *Y. lipolytica* using ssDNA oligonucleotide and achieved the editing of a key gene of lipid metabolism. Such an approach makes it possible to generate multiple mutants of this protein, hereby opening the way for future studies on the impact of residue variations on fatty acid production. Moreover, the circular RT-PCR gives a complete picture of the life cycle of sgRNA in *Y. lipolytica* and offers some understanding of their processing from 5' end maturation to degradation, providing us with keys to an optimized tool for efficient gene editing in *Y. lipolytica*, ultimately paving the way for rapid and straightforward novel strain conception in metabolic engineering and synthetic biology applications.

## Materials and Methods

### Strains and media

*E. coli* strain XL1 blue (Agilent) was used for cloning and plasmid production. For routine cloning, the cells were rendered competent using the Mix & Go Competent cells kit (Zymo Research) and transformed following the kit manufacturers protocol. Clones were grown in LB media supplemented with ampicillin (100 µg/mL) at 37 °C, and plasmids were subsequently extracted with the Plasmid Miniprep Kit (Qiagen). *Y. lipolytica* strain Y6047 (*ura3-302*, *leu2-270*, *xpr2-322* LEU2-pTEF-RedStar2) was obtained by excision from strain JMY2314 [44] of the LEU2 marker using the Cre-lox recombinase system with the replicative plasmid pUB4-Cre1 [45]. Strain Y6047 is therefore is auxotrophic for leucine and uracil, with a single copy of the RedStar2 sequence under the pTEF promoter. The

yeast was typically grown at 28–30 °C in YPD medium (1% w/v yeast extract, 1% w/v bactopectone, 1% w/v glucose) or minimum medium YNB containing yeast nitrogen base (1.7 g/L, without amino acids and ammonium sulfate), NH<sub>4</sub>Cl (5 g/L), glucose (10 g/L), and phosphate buffer (50 mM, pH 6.8), complemented when necessary with uracil (0.1 g/L) or casamino acids (2 g/L). Solid medium contains 1.5% agar.

### Target design

A target of 20 bp was designed within the RedStar2 sequence using the online tool CRISPOR-TEFOR [46], which proposed the sequence: GGTGAGGGTAGA CCATATGA and CGGGTGAAGTTCCACTGGT for RedStar2 and *FAD2*, respectively, as suitable candidates with no apparent off target site in *Y. lipolytica* genome.

### Transformation of *Y. lipolytica* with CRISPR/Cas9 plasmids

Strain Y6047 was made competent using the Frozen EZ Yeast Transformation II kit (Zymo Research). In a typical procedure with hygromycin resistance selection, aliquots of 100 µL of competent cells were transformed with 5 µg of plasmid and treated following the kit manufacturer protocol. The whole transformation mixture (~1100 µL) was plated on YPD agar medium supplemented with hygromycin B (80 µg/mL). On the other hand when using LEU2 or URA3 markers for selection, 50-µL aliquots of competent cells were transformed with 0.2 µg of plasmid, and only 30 µL of the whole transformation mixture (~550 µL) was spread on YNB medium agar plates supplemented when required with uracil or casamino acids. For HDR experiments, 4 µL of ssDNA oligonucleotide solution (100 µM) was added with the plasmid. Similarly, for the time course monitoring by FACS, (using LEU2 as a selection marker), 1.5 µg of plasmid was used with 50-µL aliquots of competent cells. The 550-µL transformation mixtures were diluted with 1 mL of YNB medium complemented with uracil and gently shaken (120 rpm) in culture tubes at 30 °C for 8 h, after which more medium was added to make a final volume of 5 mL, and the tubes were vigorously shaken (200 rpm). Over 4 days, samples from the cultures were taken at intervals and kept on ice, then directly diluted to fall within a range between  $0.5 \times 10^6$  and  $1.5 \times 10^6$  cells/mL before being incubated with DAPI at 5 µg/mL during 5 min at 21 °C and then immediately injected into a flow cytometer (MACSQuant VYB, Miltenyi Biotec). Regions for population were determined as a function of the DAPI fluorescence and DsRed fluorescence. Optimal laser and filter sets for different dyes are as follows: 405-nm laser and 450/50 BP filter for DAPI, and 561-nm laser and 615/20 BP filter for DsRed.

## HDR promoted by CRISPR/Cas9

PO1d strain was made competent using the Frozen EZ Yeast Transformation II kit (Zymo Research). Following the kit manufacturer instructions, 25- $\mu$ L aliquots of competent cells were transformed with 0.2  $\mu$ g of pCg58-FAD2 plasmid (with URA3 selection markers and *FAD2* targeting guide; see SI protocol), alongside 200 pmol of ssDNA oligonucleotide. A 50- $\mu$ L aliquot of the whole transformation mixture (~275  $\mu$ L) was spread on YNB medium agar plates supplemented with casamino acids. Colonies were picked after 3 days of outgrowth at 28 °C and grown on YPD for 20 h at 28 °C. Samples from the YPD cultures were lysed in alkaline solution to be analyzed by PCR and sequencing.

## Fluorescence/OD<sub>600 nm</sub> measurements

Colonies from CRISPR/Cas9 transformation were picked after 3 days of outgrowth on solid medium and grown in 600  $\mu$ L of YPD medium in deep well plates for 2 to 3 days, or alternatively, 200  $\mu$ L in microtiter plates for a maximum of 40 h, covered with a breathable sealing membrane (Corning), at 600 rpm, at 28 °C, in a shaking incubator (Infors HT microtron or equivalent). Samples from the cultures were collected and diluted to a  $0.6 \pm 0.4$  OD<sub>600 nm</sub>, and their fluorescence was measured on an Infinite 200 PRO (Tecan) or on a Synergy Mx (Biotek) apparatus using 554 nm for excitation wavelength and 591 nm for emission wavelength. The ratio of fluorescence over OD<sub>600 nm</sub> of each sample was then compared to the one of control cultures of the Y6047 parent strain.

## Western blotting

Three colonies per CRISPR/Cas9 transformation were grown in 5 mL of YPD medium complemented with hygromycin (200  $\mu$ g/mL) in culture tubes shaken at 200 rpm at 30 °C for 20 h. Samples of 2 mL of cultures were collected and centrifuged. To the pellet was added ~200  $\mu$ L of 0.5-mm diameter glass beads and 200  $\mu$ L of antiprotease cOmplete ULTRA EDTA-free (Roche) in sodium phosphate (pH 7.2, 50 mM) and NaCl (100 mM) solution diluted according to the manufacturer's instruction. The cells were disrupted with an MP FastPrep-24 Instrument (MP Biomedicals Inc.) by three series of vortexing (20 s, 4 m/s), with 30-s pause on ice in between each run. After centrifugation, the supernatants were collected and assayed for their protein content (Bradford reagent; BioRad). A total of 10  $\mu$ g of protein per sample was analyzed by SDS PAGE (Stain Free-any kD; BioRad) and then transferred on a nitrocellulose membrane (0.45  $\mu$ m; Protran Whatman). The membrane was blocked with skimmed milk (5% w/v) for 2 h in PBS buffer–0.5% Tween® and visualized for protein transfer using the Stain Free technology. Next, incubation with Cas9 (7A9–3A3)

Mouse monoclonal antibody (Cell Signaling Technology, Inc.) at a 1:2000 dilution took place overnight at 4 °C, followed by three subsequent washing at room temperature, and a final incubation in PBS buffer–0.5% Tween® with skimmed milk (5% w/v) containing a secondary horseradish peroxidase–conjugated goat antimouse IgG(H + L) antibody (Life Technology) at a 1:2500 dilution for 90 min. Finally, the membrane was washed three times in PBS buffer–0.5% Tween® and immersed 5 min in ECL Western Blotting Reagent (Pierce), the chemiluminescence signal was measured on a Chemidoc apparatus (BioRad) and analyzed with the Image Lab software (BioRad).

## RedStar2 locus sequencing

Samples, typically 100  $\mu$ L, of cultured colonies grown in deep well plates or microtiter plates were centrifuged. The cells were washed with physiological water and centrifuged again, and the resulting pellets were suspended in alkaline solution (NaOH 25 mM, 50  $\mu$ L) prior to being heated at 95 °C for 10 min. Supernatants could be stored at –20 °C or used directly for matrix in the PCR reactions (see SI for PCR conditions).

## cRT-PCR

Colonies from CRISPR–Cas9 transformations were grown to an OD<sub>600 nm</sub> of 1.5 in YPD medium supplemented with hygromycin (100  $\mu$ g/mL) shaking at 28 °C. For total RNA extraction, 3 mL of each culture was collected, centrifuged, washed with TE buffer, and centrifuged again. To the pellet were added 50  $\mu$ L of TE buffer, 100  $\mu$ L of 0.5 mm diameter zirconia beads (Biospec), and 200  $\mu$ L of Lysis buffer from the SurePrep small RNA purification kit (Fisher Scientific). The cells were disrupted with an MP FastPrep-24 Instrument (MP Biomedicals Inc.) by two series of vortexing (1 min, 6 m/s), with a 5-min pause on ice in between each run. After centrifugation, the RNA in the supernatants was extracted following the SurePrep small RNA purification kit manufacturer protocol. The collected RNA solutions were quantified at 260 nm using a NanoDrop ND-2000 spectrophotometer (NanoDrop Technologies, Wilmington, DE), aliquoted, and stored at –80 °C. Genomic DNA was removed from the samples using Ambion® TURBO DNA-free™ DNase (ThermoFisher). The quality of the RNAs was analyzed with the 2100 Bioanalyzer (Agilent, Palo Alto, CA) using RNA6000 Nano chips according to the manufacturer's instructions. A total of 3  $\mu$ g of RNAs was circularized with T4 RNA ligase (Takara) at 4 °C overnight in 50- $\mu$ L reactions and then purified by precipitation in NaOAc(3 M)/glycogen/EtOH. Following centrifugation, the RNA pellets were dissolved in DEPC-treated water (13  $\mu$ L) and directly reverse transcribed with the Superscript III kit (Invitrogen) using a specific primer for the sgRNA

guide (see SI Table 3). PCR with a Taq polymerase (Amersham) was performed on the resulting cDNA, and the PCR products were cloned in TOPO-TA vector (Invitrogen). After plasmid propagation in *E. coli* and subsequent extraction, the fragments were sequenced using the universal primers M13\_fw and M13\_Rev.

## Genome sequencing

*Y. lipolytica* strain Y6047 was transformed with pCg53-DsRed, pCg58-DsRed, and pCas9-URA3 plasmids, and transformants were selected on YNB plates supplemented with casamino acids. For each plasmid, two transformants were used for plasmid loss experiment. An overnight culture of a single colony of the plasmid-containing strain was grown in 10 mL of non-selective liquid YPD media at 28 °C. The next day, a subculture of this overnight preculture was grown for 24 h in the same non-selective YPD media after a 1:100 dilution. This last step was repeated three times. After three rounds of dilution and growth on non-selective YPD liquid media, the OD<sub>600 nm</sub> was measured and dilutions of the culture were performed in order to plate around 100 colonies on non-selective YPD plates. After 2 days of growth at 28 °C, colonies were replicated both on selective and on non-selective media. Two colonies were selected that were able to grow on non-selective media but did not on selective one. For DNA extraction, cells were grown at 28 °C in 10 mL of YPD medium for 2 days and then harvested by centrifugation. Cell pellets were incubated for 15 min at 37 °C in 300 µL of buffer T1 [sorbitol 1 M, EDTA 0.1 M (pH 8)] with 5 µL of zymolyase 100 T (10 mg/mL). Following centrifugation, the cell pellets were suspended in a lysis buffer [Tris 10 mM (pH 8), EDTA 1 mM, NaCl 100 mM, Triton 2%, SDS 1%] with glass beads and phenol–chloroform–isoamyl alcohol 25:24:1, vortexed for 2 min, and centrifuged. The supernatants were then precipitated with ethanol and washed twice with ethanol 70%. The resulting DNA pellets were dissolved in 100 µL TE with 1 µL RNase A (10 mg/mL), incubated at 37 °C for 15 min, and ethanol precipitated. Finally, the DNA was dissolved in Tris buffer (10 mM, pH 8). Shotgun libraries were sequenced on an Illumina HiSeq2000 sequencing platform (paired-end 2 × 125 bp) by GATC Biotech (Mulhouse, France). The resulting reads were trimmed with Trimmomatic version 0.32 [47] and cutadapt version 1.8.3 [48]. Genome assemblies were done using SOAPdenovo2 version 2.04 [49], with a k-mer estimated with kmergenie version 1.67 [50]. Superscaffolding was performed with Redundans version 0.13c [51] with W29 as a reference genome [31]. Read mapping was performed with BWA version 0.7.12 [52]. We used SAMtools version 1.2 [53], Picard tools version 2.9.0 (<http://broadinstitute.github.io/picard>), and GATK

version 3.7-0 [54] for SNPs and short indel identification and filtering. Genome alignment was performed with MUMmer 3 package [55] (nucmer version 3.1 and mummerplot version 3.5).

## Acknowledgments

The authors thank Jean-Marc Nicaud for providing the genetic elements that permitted the construction of the final plasmids of this study and François Brunel for performing excision of the *LEU2* marker from strain JMY2314, leading to strain Y6047. The authors thank the Toulouse White Biotechnology Centre for the use of the cytometry facilities and the ICEO facility part of the Integrated Screening Platform of Toulouse (PICT, IBI SA) for providing access to its screening and analytical facilities. This project was funded in the context of the call for precompetitive projects of Toulouse White Biotechnology and was co-funded by Toulouse White Biotechnology (2017/059), Région Occitanie/Pyrénées-Méditerranée (no. 5066717), and the FEDER 2014–2020 program (no. 16005006).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jmb.2018.08.024>.

Received 24 April 2018;

Received in revised form 27 July 2018;

Accepted 27 August 2018

Available online 15 September 2018

### Keywords:

oleaginous yeast;  
genome editing;  
sgRNA processing;  
sequencing

†C.N., F.D., and F.B. contributed equally to this work.

### Abbreviations used:

sgRNA, single guide RNA; DSB, double-strand break; RNP, ribonucleic acid protein; PAM, protospacer adjacent sequence; NHEJ, non-homologous end-joining; tRNA, transfer RNA; Pol III, polymerase III; KO, knockout; WT, wild-type; HDR, homology-directed repair.

## References

- [1] A. Beopoulos, J. Cescut, R. Haddouche, J.-L. Urbelarrea, C. Molina-Jouve, J.-M. Nicaud, *Yarrowia lipolytica* as a model for bio-oil production, *Prog. Lipid Res.* 48 (2009) 375–387, <https://doi.org/10.1016/j.plipres.2009.08.005>.

- [2] J.-M. Nicaud, C. Madzak, P. van den Broek, C. Gysler, P. Duboc, P. Niederberger, C. Gaillardin, Protein expression and secretion in the yeast *Yarrowia lipolytica*, *FEMS Yeast Res.* 2 (2002) 371–379, <https://doi.org/10.1111/j.1567-1364.2002.tb00106.x>.
- [3] C. Rigouin, M. Gueroult, C. Croux, G. Dubois, V. Borsenberger, S. Barbe, A. Marty, F. Daboussi, I. André, F. Bordes, Production of medium chain fatty acids by *Yarrowia lipolytica*: combining molecular design and TALEN to engineer the fatty acid synthase, *ACS Synth. Biol.* 6 (2017) 1870–1879, <https://doi.org/10.1021/acssynbio.7b00034>.
- [4] Z.-P. Guo, S. Duquesne, S. Bozonnet, G. Cioci, J.-M. Nicaud, A. Marty, M.J. O'Donohue, Conferring cellulose-degrading ability to *Yarrowia lipolytica* to facilitate a consolidated bioprocessing approach, *Biotechnol. Biofuels* 10 (2017), 132. <https://doi.org/10.1186/s13068-017-0819-8>.
- [5] K. Qiao, T.M. Wasylenko, K. Zhou, P. Xu, G. Stephanopoulos, Lipid production in *Yarrowia lipolytica* is maximized by engineering cytosolic redox metabolism, *Nat. Biotechnol.* 35 (2017) 173–177, <https://doi.org/10.1038/nbt.3763>.
- [6] J. Blazeck, A. Hill, L. Liu, R. Knight, J. Miller, A. Pan, P. Otoupal, H.S. Alper, Harnessing *Yarrowia lipolytica* lipogenesis to create a platform for lipid and biofuel production, *Nat. Commun.* 5 (2014), 3131. <https://doi.org/10.1038/ncomms4131>.
- [7] M. Groenewald, T. Boekhout, C. Neuvéglise, C. Gaillardin, P.W.M. van Dijck, M. Wyss, *Yarrowia lipolytica*: safety assessment of an oleaginous yeast with a great industrial potential, *Crit. Rev. Microbiol.* 40 (2014) 187–206, <https://doi.org/10.3109/1040841X.2013.770386>.
- [8] P. Mali, L. Yang, K.M. Esvelt, J. Aach, M. Guell, J.E. Dicarlo, J.E. Norville, G.M. Church, RNA-guided human genome engineering via Cas9, *Science* 339 (2013) 823–826, <https://doi.org/10.1126/science.1232033>.
- [9] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, E. Charpentier, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity, *Science* 337 (2012) 816–821, <https://doi.org/10.1126/science.1225829>.
- [10] P.D. Hsu, E.S. Lander, F. Zhang, Development and applications of CRISPR–Cas9 for genome engineering, *Cell* 157 (2014) 1262–1278, <https://doi.org/10.1016/j.cell.2014.05.010>.
- [11] C.M. Schwartz, M.S. Hussain, M. Blenner, I. Wheeldon, Synthetic RNA polymerase III promoters facilitate high-efficiency CRISPR–Cas9-mediated genome editing in *Yarrowia lipolytica*, *ACS Synth. Biol.* 5 (2016) 356–359, <https://doi.org/10.1021/acssynbio.5b00162>.
- [12] S. Gao, Y. Tong, Z. Wen, L. Zhu, M. Ge, D. Chen, Y. Jiang, S. Yang, Multiplex gene editing of the *Yarrowia lipolytica* genome using the CRISPR–Cas9 system, *J. Ind. Microbiol. Biotechnol.* 43 (2016) 1085–1093, <https://doi.org/10.1007/s10295-016-1789-8>.
- [13] C. Holkenbrink, M.I. Dam, K.R. Kildegaard, J. Beder, J. Dahlin, D. Doménech Belda, I. Borodina, EasyCloneYALI: CRISPR/Cas9-based synthetic toolbox for engineering of the yeast *Yarrowia lipolytica*, *Biotechnol. J. (n.d.)*. <https://doi.org/10.1002/biot.201700543>.
- [14] C. Schwartz, M. Shabbir-Hussain, K. Frogue, M. Blenner, I. Wheeldon, Standardized markerless gene integration for pathway engineering in *Yarrowia lipolytica*, *ACS Synth. Biol.* 6 (2017) 402–409, <https://doi.org/10.1021/acssynbio.6b00285>.
- [15] C. Schwartz, K. Frogue, A. Ramesh, J. Misa, I. Wheeldon, CRISPRi repression of nonhomologous end-joining for enhanced genome engineering via homologous recombination in *Yarrowia lipolytica*, *Biotechnol. Bioeng.* 114 (2017) 2896–2906, <https://doi.org/10.1002/bit.26404>.
- [16] J. Acker, C. Ozanne, R. Kachouri-Lafond, C. Gaillardin, C. Neuvéglise, C. Marck, Dicistronic tRNA–5S rRNA genes in *Yarrowia lipolytica*: an alternative TFIIIA-independent way for expression of 5S rRNA genes, *Nucleic Acids Res.* 36 (2008) 5832–5844, <https://doi.org/10.1093/nar/gkn549>.
- [17] K. Xie, B. Minkenberg, Y. Yang, Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system, *Proc. Natl. Acad. Sci.* 112 (2015) 3570–3575, <https://doi.org/10.1073/pnas.1420294112>.
- [18] C. Janke, M.M. Magiera, N. Rathfelder, C. Taxis, S. Reber, H. Maekawa, A. Moreno-Borchart, G. Doenges, E. Schwob, E. Schiebel, M. Knop, A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes, *Yeast* 21 (2004) 947–962, <https://doi.org/10.1002/yea.1142>.
- [19] P. Kabran, T. Rossignol, C. Gaillardin, J.-M. Nicaud, C. Neuvéglise, Alternative splicing regulates targeting of malate dehydrogenase in *Yarrowia lipolytica*, *DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes* 19 (2012) 231–244, <https://doi.org/10.1093/dnares/dss007>.
- [20] J.E. Dicarlo, J.E. Norville, P. Mali, X. Rios, J. Aach, G.M. Church, Genome engineering in *Saccharomyces cerevisiae* using CRISPR–Cas systems, *Nucleic Acids Res.* 41 (2013) 4336–4343, <https://doi.org/10.1093/nar/gkt135>.
- [21] J. Blazeck, L. Liu, H. Redden, H. Alper, Tuning gene expression in *Yarrowia lipolytica* by a hybrid promoter approach, *Appl. Environ. Microbiol.* 77 (2011) 7905–7914, <https://doi.org/10.1128/AEM.05763-11>.
- [22] P. Fournier, L. Guyaneux, M. Chasles, C. Gaillardin, Scarcity of ars sequences isolated in a morphogenesis mutant of the yeast *Yarrowia lipolytica*, *Yeast* 7 (1991) 25–36, <https://doi.org/10.1002/yea.320070104>.
- [23] R. Cordero Otero, C. Gaillardin, Efficient selection of hygromycin-B-resistant *Yarrowia lipolytica* transformants, *Appl. Microbiol. Biotechnol.* 46 (1996) 143–148.
- [24] E.K. Brinkman, T. Chen, M. Amendola, B. van Steensel, Easy quantitative assessment of genome editing by sequence trace decomposition, *Nucleic Acids Res.* 42 (2014) e168, <https://doi.org/10.1093/nar/gku936>.
- [25] S. Yokobori, S. Pääbo, Transfer RNA editing in land snail mitochondria, *Proc. Natl. Acad. Sci. U. S. A.* 92 (1995) 10432–10435.
- [26] M.K. Doma, R. Parker, RNA quality control in eukaryotes, *Cell* 131 (2007) 660–668, <https://doi.org/10.1016/j.cell.2007.10.041>.
- [27] Š. Vaňáčová, J. Wolf, G. Martin, D. Blank, S. Dettwiler, A. Friedlein, H. Langen, G. Keith, W. Keller, A new yeast poly(A) polymerase complex involved in RNA quality control, *PLoS Biol.* 3 (2005), e189. <https://doi.org/10.1371/journal.pbio.0030189>.
- [28] J. Lacava, J. Houseley, C. Saveanu, E. Petfalski, E. Thompson, A. Jacquier, D. Tollervey, RNA degradation by the exosome is promoted by a nuclear polyadenylation complex, *Cell* 121 (2005) 713–724, <https://doi.org/10.1016/j.cell.2005.04.029>.
- [29] M. Bühler, W. Haas, S.P. Gygi, D. Moazed, RNAi-dependent and -independent RNA turnover mechanisms contribute to heterochromatic gene silencing, *Cell* 129 (2007) 707–721, <https://doi.org/10.1016/j.cell.2007.03.038>.
- [30] S. Kadaba, A. Krueger, T. Trice, A.M. Krecic, A.G. Hinnebusch, J. Anderson, Nuclear surveillance and degradation of hypomodified initiator tRNAmet in *S. cerevisiae*, *Genes Dev.* 18 (2004) 1227–1240, <https://doi.org/10.1101/gad.1183804>.
- [31] C. Magnan, J. Yu, I. Chang, E. Jahn, Y. Kanomata, J. Wu, M. Zeller, M. Oakes, P. Baldi, S. Sandmeyer, Sequence

- assembly of *Yarrowia lipolytica* strain W29/CLIB89 shows transposable element diversity, *PLoS One* 11 (2016), e0162363. <https://doi.org/10.1371/journal.pone.0162363>.
- [32] F. Storici, J.R. Snipe, G.K. Chan, D.A. Gordenin, M.A. Resnick, Conservative repair of a chromosomal double-strand break by single-strand DNA through two steps of annealing, *Mol. Cell. Biol.* 26 (2006) 7645–7657, <https://doi.org/10.1128/MCB.00672-06>.
- [33] H. Yang, H. Wang, R. Jaenisch, Generating genetically modified mice using CRISPR/Cas-mediated genome engineering, *Nat. Protoc.* 9 (2014) 1956–1968, <https://doi.org/10.1038/nprot.2014.134>.
- [34] A. Beopoulos, J. Verbeke, F. Bordes, M. Guicherd, M. Bressy, A. Marty, J.-M. Nicaud, Metabolic engineering for ricinoleic acid production in the oleaginous yeast *Yarrowia lipolytica*, *Appl. Microbiol. Biotechnol.* 98 (2014) 251–262, <https://doi.org/10.1007/s00253-013-5295-x>.
- [35] G. Barth, C. Gaillardin, *Yarrowia lipolytica*, *Nonconv. Yeasts Biotechnol.*, Springer, Berlin, Heidelberg 1996, pp. 313–388, [https://doi.org/10.1007/978-3-642-79856-6\\_10](https://doi.org/10.1007/978-3-642-79856-6_10).
- [36] J.-B. Renaud, C. Boix, M. Charpentier, A. De Cian, J. Cochennec, E. Duvernois-Berthet, L. Perrouault, L. Tesson, J. Edouard, R. Thinar, Y. Cherifi, S. Menoret, S. Fontanière, N. de Crozé, A. Fraichard, F. Sohm, I. Anegon, J.-P. Concordet, C. Giovannangeli, Improved genome editing efficiency and flexibility using modified oligonucleotides with TALEN and CRISPR–Cas9 nucleases, *Cell Rep.* 14 (2016) 2263–2272, <https://doi.org/10.1016/j.celrep.2016.02.018>.
- [37] E.K. Brinkman, A.N. Kousholt, T. Harmsen, C. Leemans, T. Chen, J. Jonkers, B. van Steensel, Easy quantification of template-directed CRISPR/Cas9 editing, *Nucleic Acids Res.* 46 (2018) e58, <https://doi.org/10.1093/nar/gky164>.
- [38] J.Z. Jacobs, K.M. Ciccaglione, V. Tournier, M. Zaratiegui, Implementation of the CRISPR–Cas9 system in fission yeast, *Nat. Commun.* 5 (2014), 5344. <https://doi.org/10.1038/ncomms6344>.
- [39] O.W. Ryan, J.M. Skerker, M.J. Maurer, X. Li, J.C. Tsai, S. Poddar, M.E. Lee, W. Deloache, J.E. Dueber, A.P. Arkin, J.H. Cate, Selection of chromosomal DNA libraries using a multiplex CRISPR system, *elife* 3 (2014), e03703. <https://doi.org/10.7554/eLife.03703>.
- [40] H.G. Damude, H. Zhang, L. Farrall, K.G. Ripp, J.-F. Tomb, D. Hollerbach, N.S. Yadav, Identification of bifunctional  $\Delta 12/\omega 3$  fatty acid desaturases for improving the ratio of  $\omega 3$  to  $\omega 6$  fatty acids in microbes and plants, *Proc. Natl. Acad. Sci.* 103 (2006) 9446–9451, <https://doi.org/10.1073/pnas.0511079103>.
- [41] A. Buček, P. Matoušková, H. Sychrová, I. Pichová, O. Hrušková-Heidingsfeldová,  $\Delta 12$ -fatty acid desaturase from *Candida parapsilosis* is a multifunctional desaturase producing a range of polyunsaturated and hydroxylated fatty acids, *PLoS One* 9 (2014), e93322. <https://doi.org/10.1371/journal.pone.0093322>.
- [42] Y. Fu, J.D. Sander, D. Reyon, V.M. Cascio, J.K. Joung, Improving CRISPR–Cas nuclease specificity using truncated guide RNAs, *Nat. Biotechnol.* 32 (2014) 279–284, <https://doi.org/10.1038/nbt.2808>.
- [43] J. Houseley, D. Tollervey, Yeast Trf5p is a nuclear poly(A) polymerase, *EMBO Rep.* 7 (2006) 205–211, <https://doi.org/10.1038/sj.embor.7400612>.
- [44] Z. Lazar, T. Rossignol, J. Verbeke, A.-M. Crutz-Le Coq, J.-M. Nicaud, M. Robak, Optimized invertase expression and secretion cassette for improving *Yarrowia lipolytica* growth on sucrose for industrial applications, *J. Ind. Microbiol. Biotechnol.* 40 (2013) 1273–1283, <https://doi.org/10.1007/s10295-013-1323-1>.
- [45] P. Fickers, M.T. Le Dall, C. Gaillardin, P. Thonart, J.M. Nicaud, New disruption cassettes for rapid gene disruption and marker rescue in the yeast *Yarrowia lipolytica*, *J. Microbiol. Methods* 55 (2003) 727–737.
- [46] M. Haeussler, K. Schönig, H. Eckert, A. Eschstruth, J. Mianné, J.-B. Renaud, S. Schneider-Maunoury, A. Shkumatava, L. Teboul, J. Kent, J.-S. Joly, J.-P. Concordet, Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR, *Genome Biol.* 17 (2016), 148. <https://doi.org/10.1186/s13059-016-1012-2>.
- [47] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, *Bioinformatics* 30 (2014) 2114–2120, <https://doi.org/10.1093/bioinformatics/btu170>.
- [48] M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, *EMBnet J.* 17 (2011) 10–12, <https://doi.org/10.14806/ej.17.1.200>.
- [49] R. Luo, B. Liu, Y. Xie, Z. Li, W. Huang, J. Yuan, G. He, Y. Chen, Q. Pan, Y. Liu, J. Tang, G. Wu, H. Zhang, Y. Shi, Y. Liu, C. Yu, B. Wang, Y. Lu, C. Han, D.W. Cheung, S.-M. Yiu, S. Peng, Z. Xiaoqian, G. Liu, X. Liao, Y. Li, H. Yang, J. Wang, T.-W. Lam, J. Wang, SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler, *GigaScience* 1 (2012) 1–6, <https://doi.org/10.1186/2047-217X-1-18>.
- [50] R. Chikhi, P. Medvedev, Informed and automated k-mer size selection for genome assembly, *Bioinformatics* 30 (2014) 31–37, <https://doi.org/10.1093/bioinformatics/btt310>.
- [51] L.P. Pryszyk, T. Gabaldón, Redundans: an assembly pipeline for highly heterozygous genomes, *Nucleic Acids Res.* 44 (2016) e113, <https://doi.org/10.1093/nar/gkw294>.
- [52] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform, *Bioinformatics* 25 (2009) 1754–1760, <https://doi.org/10.1093/bioinformatics/btp324>.
- [53] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The sequence alignment/map format and SAMtools, *Bioinformatics* 25 (2009) 2078–2079, <https://doi.org/10.1093/bioinformatics/btp352>.
- [54] A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytzky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M.A. DePristo, The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data, *Genome Res.* 20 (2010) 1297–1303, <https://doi.org/10.1101/gr.107524.110>.
- [55] S. Kurtz, A. Phillippy, A.L. Delcher, M. Smoot, M. Shumway, C. Antonescu, S.L. Salzberg, Versatile and open software for comparing large genomes, *Genome Biol.* 5 (2004) R12, <https://doi.org/10.1186/gb-2004-5-2-r12>.