

SCIENTIFIC REPORTS



OPEN

EUF1 – a newly identified gene involved in erythritol utilization in *Yarrowia lipolytica*

Dorota A. Rzechonek¹, Cécile Neuvéglise², Hugo Devillers², Waldemar Rymowicz¹ & Aleksandra M. Mironczuk¹

The gene *YAL10F01562g* was identified as an important factor involved in erythritol catabolism of the unconventional yeast *Yarrowia lipolytica*. Its putative role was identified for the first time by comparative analysis of four *Y. lipolytica* strains: A-101.1.31, *Wratislavia* K1, MK1 and AMM. The presence of a mutation that seriously damaged the gene corresponded to inability of the strain *Wratislavia* K1 to utilize erythritol. RT-PCR analysis of the strain MK1 demonstrated a significant increase in *YAL10F01562g* expression during growth on erythritol. Further studies involving deletion and overexpression of the selected gene showed that it is indeed essential for efficient erythritol assimilation. The deletion strain *Y. lipolytica* AMM Δ euf1 was almost unable to grow on erythritol as the sole carbon source. When the strain was applied in the process of erythritol production from glycerol, the amount of erythritol remained constant after reaching the maximal concentration. Analysis of the *YAL10F01562g* gene sequence revealed the presence of domains characteristic for transcription factors. Therefore we suggest naming the studied gene *Erythritol Utilization Factor* – *EUF1*.

Yarrowia lipolytica is an unconventional yeast with high potential for use in industry. One of the most interesting metabolites is erythritol, a four-carbon polyol, produced under hyperosmotic conditions. Erythritol gained interest because of its application as a sweetener, which not only provides no calories to the human body, but may also prevent development of caries. It occurs naturally in small amounts in foods such as honey, dairy products and fruits. For the industrial scale it is produced by yeasts such as *Moniliella pollinis* and *Torula* sp., mostly from glucose¹. *Y. lipolytica* can synthesize erythritol from glucose, though the preferred substrate is glycerol². Moreover, it can utilize crude glycerol without any prior purification³. The conditions necessary for efficient production are low pH and high osmotic pressure of the environment, induced by high concentrations of substrates or salt⁴. An impediment to this process is the ability of *Y. lipolytica* to utilize erythritol as a carbon source. Once the initial substrate is depleted, concentrations of secondary metabolites, including erythritol, begin to fall rapidly. In the case of batch reactors, the delay in termination of culture can result in a significant decrease in productivity.

Erythritol catabolism has been subjected to the most in-depth research in the case of the bacteria *Brucella* spp., where erythritol utilization is considered responsible for triggering the virulence that leads to fetus abortion of farm animals⁵. The genes organized in the *EryABCD* operon encode erythritol kinase (EryA), which converts erythritol to L-erythritol-4-phosphate⁶, two putative dehydrogenases (EryB and EryC) and a repressor (EryD)⁷. The knowledge of erythritol utilization in yeast is much poorer, although some studies have been conducted for *Lipomyces starkeyi*⁸. The opportunity to approach the problem from another perspective came from observations of phenotypes of a few *Yarrowia lipolytica* strains from the Department of Biotechnology and Food Microbiology at Wrocław University of Environmental and Life Sciences (Poland). The wild strain *Y. lipolytica* A101 was isolated from soil⁹. The A-101 derivative strains 1.31, *Wratislavia* K1 and MK1 were obtained as a result of a series of UV and spontaneous mutations (Fig. 1). The most interesting phenotype was observed in *Wratislavia* K1 (hereinafter referred to as K1), which was created as a result of spontaneous mutation of strain 1.31. The mutation occurred during continuous citric acid production from glucose in a nitrogen-limited chemostat¹⁰. In contrast to the parental strain 1.31, which was used mainly for citric acid production, K1 was a better erythritol producer¹¹. The most characteristic feature of the latter strain was its inability to utilize erythritol once produced, while grown

¹Department of Biotechnology and Food Microbiology, Wrocław University of Environmental and Life Sciences, Chelmonskiego 37, Wrocław, 57-630, Poland. ²Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, Paris, France. Correspondence and requests for materials should be addressed to A.M.M. (email: aleksandra.mironczuk@upwr.edu.pl)

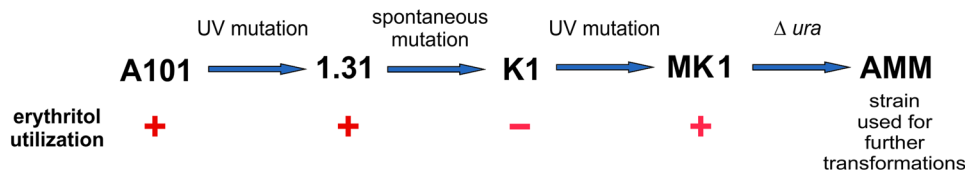


Figure 1. Background of *Yarrowia lipolytica* strains used in the study.

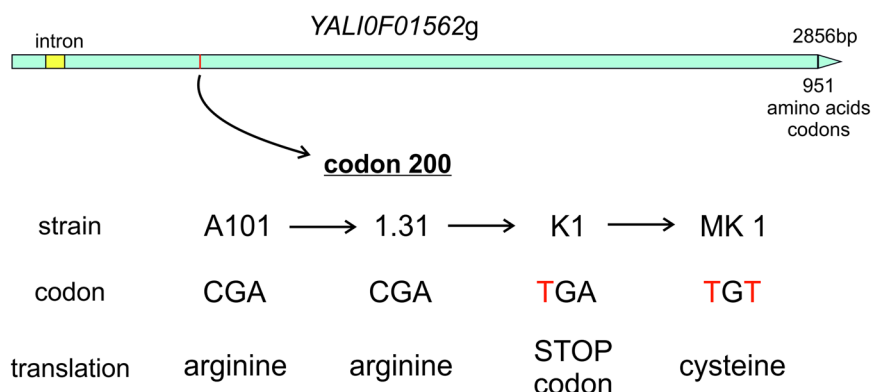


Figure 2. Successive mutations in gene *YAL10F01562g*.

on an erythrithol synthesis medium. K1 was subjected to UV mutagenesis to further improve production parameters. The resulting strain MK1 is characterized by very low by-product formation, but it also recovered the ability to utilize erythrithol¹². The interesting properties of strains 1.31, K1 and MK1 were the rationale for full genome sequencing and mapping.

The aim of this study was to identify genes responsible for erythrithol utilization in *Y. lipolytica*.

Results

Mutations in gene *YAL10F01562g*. Comparative analysis of *Y. lipolytica* strains 1.31, K1 and MK1 genomes against A101 scaffolds indicated the existence of 18 mutations between strains 1.31 and K1 and an additional 13 mutations between K1 and MK1. Particularly interesting progression of changes was noted in the gene *YAL101S02e22540*, a homolog of *YAL10F01562g* in the genome of *Yarrowia lipolytica* strain CLIB122. Because of the greater dissemination of CLIB122 annotation, its gene name will be used in the following description of the results. *YAL10F01562g* encodes a protein of 951 amino acids. The gene carries at the 5' end a small intron of 78 bp. A first mutation appeared in strain K1. In the first position of codon 200 there was a transition of cytosine to thymine that changed the arginine codon CGA to a terminal codon TGA (stop). In strain MK1 this mutation was also present. However, in the third position of the same codon, another point mutation occurred. An adenine to thymine transversion changed the stop codon into a cysteine codon, TGT (Fig. 2). With the restoration of the gene, strain MK1 recovered the ability to utilize erythrithol. Thus, we decided to further investigate *YAL10F01562g*.

Comparative analysis with yeast databases. Comparing *YAL10F01562g* translated into the protein sequence with the NCBI databases via the BLASTp algorithm showed that the protein contains sequences resembling two putative conserved domains: a GAL-4 like binuclear cluster DNA-binding domain at the beginning of the sequence (approximately codons 20–65) and a fungal specific transcription factor domain (codons 279–560)^{13,14}. Search for a homology to whole known proteins showed some similarity to the C6 transcriptional factor from *Rhodotorula toruloides*, and to a pathway-specific nitrogen regulator from *Moniliophthora roreri* and *Cryptococcus neoformans*. However, the highest similarity was recorded for hypothetical proteins from *Lipomyces starkeyi* NRRL Y-11557 (Table S1). Because of these results we preliminarily described the gene as *EUF1*, standing for Erythrithol Utilization Factor, and this name was used in the nomenclature of modified strains.

Transcript level studies. The function of the gene *YAL10F01562g* was first investigated by measuring its transcript level on different carbon sources: glucose, glycerol or erythrithol. Samples for RNA isolation were collected in 24-h cultures. The references for qRT-PCR were samples from glucose cultures. Growth on glycerol did not modify the expression of *YAL10F01562g*; transcripts for glycerol and glucose samples were identical (Fig. 3). In contrast, when erythrithol was used as a sole carbon source, the expression of *YAL10F01562g* rose tenfold. These results showed that *YAL10F01562g* plays a role in erythrithol metabolism, since its expression was significantly enhanced in the presence of this polyol. Therefore, the next step was an investigation of its influence on *Y. lipolytica* phenotype.

Growth on different carbon sources. The next stages of the work were deletion, overexpression and complementation of the *YAL10F01562g* gene. Transformants of *Y. lipolytica* for deletion (*AMM Δeuf1*),

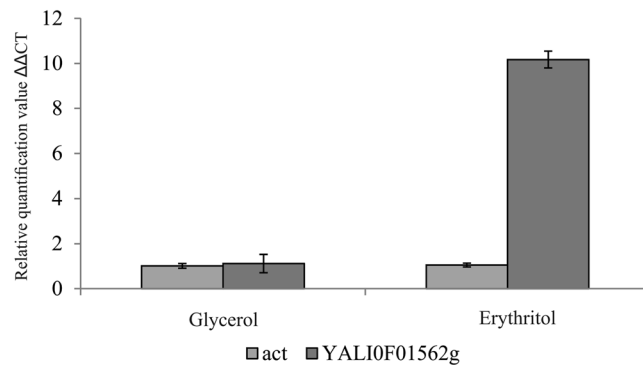


Figure 3. Relative expression of *YALI0F01562g* on media with glycerol or erythritol as a sole carbon source, with glucose medium used as a reference. Actin was used as a reference gene. Samples were analysed in triplicate, and the standard errors were estimated using Illumina Eco software.

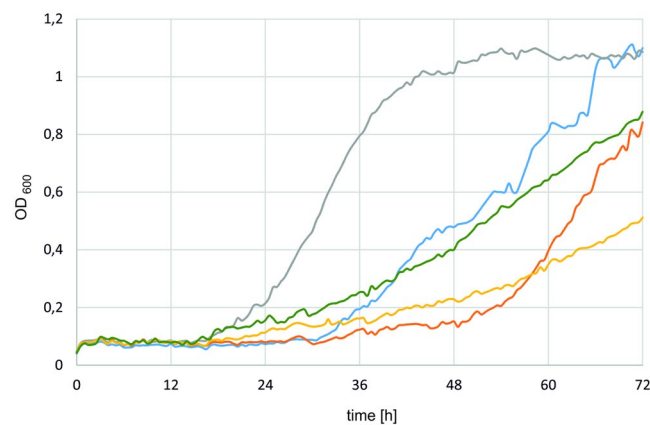


Figure 4. Growth curves of *Y. lipolytica* strains: MK1 (blue), K1 (orange), AMM Δ euf1 (yellow), AMM pAD-euf1 (grey) and AMM C-euf1 (green). The strains were grown on YNB supplemented with 5% erythritol as the sole carbon source. Experiment was performed at 28 °C under constant agitation using Bioscreen C.

overexpression (AMM pAD-euf1) and complementation (AMM C-euf1) were tested for their ability to grow on different carbon sources, i.e. glucose, erythritol and glycerol using Bioscreen C. In the course of transformations the strains regain prototrophy for uracil; thus their growth can be compared. Strains MK1 and K1 were used as controls. As seen in Fig. 4, differences in growth rate were observed when erythritol was used as the sole carbon source. However, the significance of these changes was strongly influenced by the method of inoculum preparation. *Y. lipolytica* is known for lipid accumulation, so cells used as inoculum were cultivated in medium providing only a minimal amount of nutrients (YNB). If the inoculum cells were grown in rich medium such as YPD the growth rate of all strains was similar, probably due to the use of accumulated back-up materials (data not shown). Therefore, to ensure that observed growth was a result of only erythritol assimilation, the inoculum was prepared for 7 days (two cultures lasting 48 hours and one 72 hours in YNB + 2% glucose). After such treatment all strains had a long lag phase. The growth of strain AMM pAD-euf1 started at around 18 hours, and it was the only strain that reached the stationary phase by the end of the experiment. The growth of MK1 started 12 hours later, and its lag phase was longer than all modified strains, but it finally reached the highest biomass production. The slowest growth was observed for the K1 strain, which had a lag phase of more than 48 hours. However, after that time exponential growth occurred. This result was unexpected as K1 was previously believed not to utilize erythritol. The growth of the AMM Δ euf1 strain was also impaired, though the pattern was different. The growth started after 18 hours and was very slow but linear (Fig. 4).

The complementation strain AMM C-euf1 with a copy of the *YALI0F01562g* gene under the UAS1B₁₆-TEF promoter displayed improved growth, but despite the shorter lag phase it reached a lower OD₆₀₀ value than MK1. Similar experiments were performed for glucose and glycerol as the sole carbon sources, but there were no significant changes in growth of the tested strains (data not shown).

Utilization of erythritol. The next step was to determine the erythritol utilization ratio. The initial erythritol concentration was 100 g/L in a working volume of 50 mL. During the first 24 hours of the experiment the strain AMM pAD-euf1 exhibited a significantly larger decrease in erythritol concentration (Fig. 5) than the other strains. This 24-hour advantage over the MK1 strain remained visible until complete depletion of the substrate.

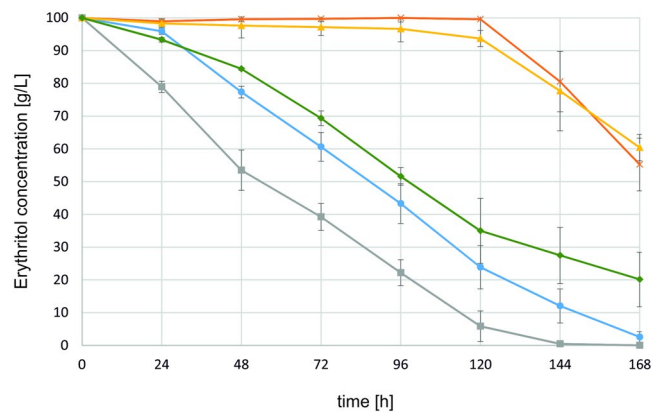


Figure 5. Utilisation of erythritol by *Y. lipolytica* strains MK1 (blue), K1 (orange), AMM Δ euf1 (yellow), AMM pAD-euf1 (grey) and AMM C-euf1 (green). The strains were grown in YNB medium supplemented with 10% erythritol.

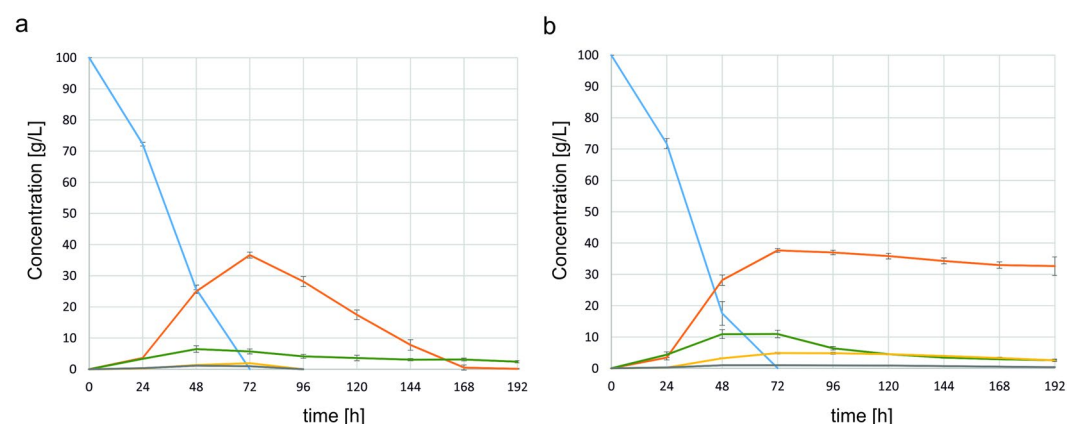


Figure 6. Assimilation of glycerol (blue), production of erythritol (orange), citric acid (green) and mannitol (yellow) and arabitol (grey) by strains MK1 (A) and AMM Δ euf1 (B) during the growth on erythritol production medium.

The complementation strain AMM c-euf1 displayed a slightly worse utilization ratio than MK1, even though it was slightly more effective during the first 24 hours.

Strains K1 and AMM Δ euf1 were also able to utilize erythritol, but the utilization of the substrate occurred only after 120 hours. Very unusual growth was observed in AMM Δ euf1 cultures; namely, the glass walls of the flasks were covered by a thick layer of biomass, while the medium beneath remained clear (Supplementary data). The efficient utilization started after the biomass layer was shaken off into liquid medium. Biofilm formation was observed also in K1 cultures, but to a smaller degree. In strains capable of fast erythritol utilization, cell growth occurred only in the liquid medium.

Finally, the impact of *YALI0F01562g* deletion on erythritol production was investigated during a shake-flask experiment on the Erythritol Synthesis Medium (ESM medium) (Fig. 6). The first 72 hours of culture proceeded similarly for both *Y. lipolytica* AMM Δ euf1 and MK1. By this time, glycerol, the main carbon source, was already depleted and the concentration of erythritol reached its maximum value of 37.6 ± 0.6 g/L for AMM Δ euf1 and 36.6 ± 1.0 g/L for MK1. In both cases mannitol and citric acid were formed as by-products; their concentration was higher for the AMM Δ euf1 strain. At the 96th hour, the concentration of erythritol in the MK1 culture decreased to 28.1 ± 1.6 g/L, and by the 168th hour all the polyols were utilized. In AMM Δ euf1 culture the amount of erythritol remained high, i.e. 37 ± 0.8 g/L at 96 h and 32.6 ± 3.0 g/L at 192 h. Mannitol was also still present, in the concentration of 2.5 g/L. The final amount of citric acid, 2.6 g/L, was similar to that in MK1. In both cultures a small amount of arabitol was detected (in a maximal concentration of 1.14 g/L). It was quickly utilized by MK1, but remained in AMM Δ euf1 at a final concentration of 0.37 g/L.

The maximal erythritol productivity of the MK1 strain was 0.51 g/L/h with yield 0.37 g/g after 72 hours of culture. When the experiment was prolonged to 96 hours, the yield value decreased to 0.28 g/g. The maximal productivity of the AMM pAD-euf1 strain was 0.52 g/L/h and yield 0.38 g/g after 72 hours of culture. The extension of culture to 96 h did not change the yield of erythritol.

Discussion

Until now no genes involved in erythritol catabolism in yeasts were known. The identification of *YALIOF01562g* was possible through comparative genomic analysis and the occurrence of mutations in strains K1 and MK1. The deletion of the gene in a *ura*⁻ derivative of MK1 led to formation of a strain with phenotypic features more similar to K1. Strains AMM Δ euf1 and K1 share the feature of delayed induction of erythritol utilization, which significantly impaired growth on erythritol as a sole carbon source (Figs 4, 5).

The complementation strain AMM C-euf1, where *YALIOF01562g* was first deleted from its native locus and later introduced under the UAS1B₁₆-TEF promoter, attested that impaired erythritol utilization was a result of *YALIOF01562g* deletion. After complementation the ability to grow and utilize erythritol was recovered, but it was slightly worsened compared to MK1 (Figs 4, 5). This may be a result of weakening of the strain by a series of transformations (*YALIOF01562g* deletion, resorting of auxotrophy and insertion of the UAS1B₁₆-TEF *YALIOF01562g* cassette).

The experiments presented in this study demonstrate the importance of *YALIOF01562g* for erythritol catabolism. However, its exact role is still unclear. Analysis of the primary structure of the encoded protein indicated the presence of a DNA-binding domain and fungal specific transcription factor domain. It strongly suggests that *YALIOF01562g* may be a transcription factor rather than an enzyme directly involved in erythritol transformations, which is also indirectly implied by the obtained results.

The proteins responsible for erythritol catabolism are unlikely to be consistently present in yeast cells. This is suggested by the relatively long lag phase of cultures on erythritol utilization medium (Fig. 5). After 24 hours of MK1 growth, only a slight change in erythritol concentrations was observed. At the same time, the MK1 strain grown on glycerol or glucose as the sole carbon sources is able to utilize up to 25% of the initial substrate (data not shown). This delay may be caused by the necessity of synthesizing the required enzymes, and, if so, their formation must be induced by some intracellular signal. Thus, the disorder of the regulation system may result in the observed phenotypes of K1 and AMM Δ euf1 strains. Initially, the K1 strain was considered as unable to catabolize erythritol. However, further research demonstrated that induction of the process is possible, but with a significant delay compared to strains with a functional *YALIOF01562g* gene.

If there was an enzyme deletion, the metabolic pathway should be blocked. Some enzymes have a few orthologs that catalyze the same reaction, but if that was the reason, the existence of the orthologous gene should have been detected during BLAST analysis.

We would also like to underline the unusual growth of the AMM Δ euf1 strain in the shake flask cultures when erythritol was applied as the sole carbon source. Initially the cells were able to grow in biofilm on the vessel walls, but not in the liquid medium. Only after this step did efficient erythritol utilization occur. Biofilm growth in yeasts is associated with presence of specific transcriptional factors, such as *Bcr1* in *Candida albicans*¹⁵. Their activity leads to significant morphological and physiological changes of the cells. Thus, we suspect that an unknown factor, other than *YALIOF01562g*, may influence erythritol catabolism, but its activation requires specific conditions or prolonged time.

The suggestion that *YALIOF01562g* might have broader influence than erythritol catabolism alone comes from the observation of AMM Δ euf1 growth on erythritol production medium (Fig. 6b). At the end of the experiment, aside from erythritol, there were two other polyols present in the supernatant: mannitol and arabitol. All these compounds were completely depleted by the MK1 strain. These results imply that *YALIOF01562g* should be further investigated for its role in catabolism regulation of the whole polyol group. The deletion did not have an influence on glycerol assimilation, but the catabolism of this polyol might undergo more complicated control, while it involves the enzymes of the glycolysis pathway¹⁶.

The deletion of *YALIOF01562g* induced the occurrence of AMM Δ euf1 phenotype resembling K1. However, on erythritol utilization medium the strains displayed different patterns of growth. The main reason is still unknown, although there are some possible explanations. First there were more mutations between K1 and MK1 strains, and their influence has not been examined yet. Secondly, the stop codon in the *YALIOF01562g* gene of K1 appeared at position 200, after the DNA binding site.

These data provide an insight into the role of *YALIOF01562g* in *Y. lipolytica* metabolism, and due to these interesting results the research is still in progress. Moreover, the obtained data are already important for industrial applications. The deletion strain AMM Δ euf1 synthesizes erythritol comparatively to MK1, but the defect in its simultaneous utilization may lead to higher erythritol concentration. Moreover, the deletion of *YALIOF01562g* may be combined with other genetic modifications, such as overexpression of genes involved in glycerol assimilation¹⁶, to further improve production parameters.

Conclusion

Our study is the first work dedicated to the regulation of erythritol utilization in yeasts. A putative function of the *YALIOF01562g* gene in erythritol catabolism was proposed. However, due to the high complexity of regulation pathways, the determination of its full role requires further studies. Thus, we would like to suggest the name *EUF1* – for erythritol utilization factor. We believe that our work will allow for a better understanding of the metabolic regulation of erythritol, not only in *Yarrowia lipolytica* but more broadly in other yeasts. *EUF1* shows some similarity to a number of proteins present in other yeasts, whose role has so far remained unknown or putative. Occurrence of the *EUF1* homolog in the only species of yeast which was previously tested for erythritol utilization indicates that its function may be universal in a large group of fungi. The study should also undoubtedly influence the optimization of erythritol production by *Yarrowia lipolytica*.

Methods

Microorganisms. Strains used in this study were *Y. lipolytica* A1.31, K1, MK1 and AMM, which is a *Aura3* derivative of MK1 (Table S2). These strains belong to the Department of Biotechnology and Food Microbiology at Wrocław University of Environmental and Life Sciences, Poland.

Media and culture conditions. *Escherichia coli* strains were cultivated in LB medium according to standard protocols¹⁷. Rich yeast extract peptone glucose (YPD) medium, containing 1% (w/v) yeast extract, 1% (w/v) peptone and 2% (w/v) glucose, was used to obtain yeast biomass for DNA extraction and inoculum preparation. Medium containing YNB without amino acids (Sigma-Aldrich) supplied with 2% (w/v) glucose was used for yeast inoculum preparation.

During shake-flask experiments the cultures were grown in 0.3 L baffled flasks containing 0.03 L or 0.05 L medium on a rotary shaker (CERTOMAT IS, Sartorius Stedim Biotech) at 28 °C and 240 rpm. Erythritol utilization rate was examined on medium containing YNB and 5% (w/v) or 10% (w/v) erythritol (YNB-e medium). Erythritol synthesis was conducted in Erythritol Synthesis Medium (ESM medium) containing 100 g/L glycerol (Chempur, Poland), 2.3 g/L (NH₄)₂SO₄ (Chempur), 1 g/L MgSO₄ × 7H₂O (Chempur), 0.23 g/L KH₂PO₄ (Chempur), 26.4 g/L NaCl (Chempur), 1 g/L yeast extract (Merck, Germany) and 3 g/L CaCO₃, pH 3.0. CaCO₃ was added separately to each flask after establishing pH 3 in order to prevent a fall of pH value.

Bioscreen C. The yeast strains were grown in 100-well plates in 150 μL of YNB supplemented with glucose 5% (w/v), erythritol 5% (w/v) or glycerol 5% (w/v). First, the strains were grown for 24, 48 or 72 h or in YNB medium containing 2% (w/v) glucose. The inoculum was grown two or three times on YNB medium to ensure that yeast cells did not accumulate nutrients. Finally the cells were inoculated to an OD₆₀₀ value of 0.15 in each well. Quintuple experiments were performed at 28 °C under constant agitation with a Bioscreen C system (Oy Growth Curves Ab Ltd., Finland). Growth was monitored by measuring the optical density at 420–560 nm every 30 min for 72 h.

Sequencing. Genomic DNA was extracted with the EURX Bacterial & Yeast Genomic DNA Purification Kit (EurX, Poland) from overnight cultures of selected strains, grown on YPD medium. Sequencing was performed with the Illumina MiSeq DNA sequencing platform (paired-end [PE] 2 × 250 bp).

Gene mapping and SNP analysis. The raw reads of each sequenced strain were trimmed with Trimmomatic version 0.32¹⁸ and cutadapt version 1.8.3¹⁹. The clean reads were mapped with BWA version 0.7.12²⁰ against the *Yarrowia lipolytica* A101 reference genome²¹. Single nucleotide polymorphism (SNP) or insertions and deletions (Indels) were identified on the basis of the mpileup files generated by SAMtools version 1.2²². The position of the SNPs and INDELS within the chromosome was visualized with Artemis (version 16.0.0) in order to select mutations located inside a coding DNA sequence (CDS). Candidate CDSs and their protein transcripts were compared via BLAST algorithms on the National Centre of Biotechnology Information (NCBI) and Genome Resources for Yeast Chromosomes (GRYC) websites.

RNA isolation and qRT-PCR. The shake flask cultures were grown for 24 hours in YNB medium supplemented with a 5% (w/v) carbon source, i.e. glycerol, glucose or erythritol. Next, the cultures were collected and centrifuged for 5 min at 12,000 g. The RNA was extracted using the Total RNA Mini Plus kit (A&A Biotechnology, Poland). Each sample was treated with DNase I (Thermo Scientific) according to the manufacturer's instructions. RNA quantities were measured using a Biochrom WPA Biowave II spectrophotometer (Biochrom Ltd., UK) equipped with a TrayCell (Hellma Analytics, Germany), and the samples were stored at –80 °C. Synthesis of cDNA was conducted using Maxima First Strand cDNA. Synthesis kits for RT-qPCR (Thermo Scientific) were used according to the manufacturer's instructions. qRT-PCR analyses were carried out using a DyNAmo Flash SYBR Green qPCR Kit (Thermo Scientific) and the Eco Real-Time PCR System (Illumina, USA).

Primers for RT-PCR were designed for the gene *YAL10F01562g*, which has one intron at the beginning of the sequence. Primers qF01562-F and qF01562-R bind to the first (nt 28 in the gene sequence) and second exon (nt 283), respectively, resulting in a 178-bp qRT-PCR product. The results were normalized to the actin gene *YAL10D08272g* amplified with primers ACT-F/ACT-R and analyzed using the ddCT method²³. Samples were analyzed in triplicate.

Cloning and transformation protocols. All restriction enzymes were purchased from FastDigest Thermo Scientific and all of the digestions were performed according to standard protocols. The PCR reactions were set up using recommended conditions and Phusion high-fidelity DNA polymerase (Thermo Scientific). The ligation reactions were performed for 10 min at room temperature using T4 DNA Ligase (Thermo Scientific). The gel extractions were performed using the Gel Out extraction kit purchased from A&A Biotechnology (Poland). The *E. coli* minipreps were performed using the Plasmid Mini Kit (A&A Biotechnology). Transformation of *E. coli* strains was performed using standard chemical protocols¹⁷.

For transformation of *Yarrowia lipolytica* only strains with auxotrophy for uracil were used.

Transformation was performed according to the lithium acetate method²⁴ and transformants were plated out on selective media without uracil. They were analyzed for proper integration by gDNA extraction and PCR amplification with two primer pairs. Genomic DNA (gDNA) was extracted from *Y. lipolytica* using the Genomic Mini AX Yeast Spin kit (A&A Biotechnology, Poland).

Construction of *EUF1* deletion cassette. First, the plasmid pUC-ura for deletion was created. Primers ura-PmeI-F and ura-PmlI-R amplified the lox1-Ura-lox4 region from plasmid JM1133²⁵. Next, the obtained PCR (1417 bp) product was phosphorylated and cloned into pUC18 digested with *SmaI*, resulting in the

pUC_ura vector. Next, the *YAL10F01562g* promoter region and terminator were amplified by PCR using primers pF01562-HindIII-F and pF01562-Sall-R for the promoter region and tF01562-NotI-F and tF01562-PmeI-R for the terminator region. The PCR promoter fragment (1041 bp) was digested with *HindIII* and *Sall*, and cloned into corresponding sites of pUC_ura, yielding the pUC-ura-pF01562 vector. Then, the PCR terminator fragment was digested with *PmeI* and *NotI*, and cloned into corresponding sites of pUC-ura-pF01562, resulting in the pUC-ura- Δ F01562 vector. The obtained vector was digested with *HindIII* and *PmeI* and transformed into *Y. lipolytica* AMM²⁶, obtaining the *Y. lipolytica* AMM Δ euf1 strain. Proper integration was verified by gDNA extraction and PCR analysis. Sequences of all primers used in the study are listed in Supplementary Table S3.

Auxotrophy of strain AMM Δ euf1 was restored via excision using the Cre-lox recombinase system following transformation with the replicative plasmid pUB4-Cre1 (JME547).

Construction of overexpression cassette. The gene *YAL10F01562g* was amplified from *Y. lipolytica* DNA with primers F01562-AscI-F and F01562-NheI-R, resulting in a 3030 bp PCR fragment. It was digested with the enzymes *SgsI* and *NheI* and cloned into corresponding sites of plasmid pAD²⁷, carrying the UAS1B₁₆-TEF promoter. The obtained plasmid pAD-F01562 was digested with *MssI* to create linear expression cassettes devoid of *E. coli* DNA and surrounded by *Y. lipolytica* rDNA for targeted integrations. It was used to transform *Y. lipolytica* AMM or *Y. lipolytica* AMM Δ F01562 ura-, resulting in strains AMM pAD-euf1 and AMM C-euf1, respectively.

References

- Moon, H. J., Jeya, M., Kim, I. W. & Lee, J. K. Biotechnological production of erythritol and its applications. *Appl Microbiol Biotechnol* **86**, 1017–1025 (2010).
- Rymowicz, W., Rywinska, A. & Marcinkiewicz, M. High-yield production of erythritol from raw glycerol in fed-batch cultures of *Yarrowia lipolytica*. *Biotechnol Lett* **31**, 377–380 (2009).
- Dobrowolski, A., Mitula, P., Rymowicz, W. & Mirończuk, A. M. Efficient conversion of crude glycerol from various industrial wastes into single cell oil by yeast *Yarrowia lipolytica*. *Bioresour Technol* **207**, 237–243 (2016).
- Liu, H.-H., Ji, X.-J. & Huang, H. Biotechnological applications of *Yarrowia lipolytica*: Past, present and future. *Biotechnology Advances*. <https://doi.org/10.1016/j.biotechadv.2015.07.010> (2015).
- Petersen, E. *et al.* Erythritol triggers expression of virulence traits in *Brucella melitensis*. *Microbes Infect* **15**, 440–449 (2013).
- Lillo, A. M., Tetzlaff, C. N., Sangari, F. J. & Cane, D. E. Functional expression and characterization of EryA, the erythritol kinase of *Brucella abortus*, and enzymatic synthesis of L-erythritol-4-phosphate. *Bioorganic & medicinal chemistry letters* **13**, 737–739 (2003).
- Rodriguez, M. C. *et al.* Evaluation of the effects of erythritol on gene expression in *Brucella abortus*. *PLoS One* **7**, e50876, <https://doi.org/10.1371/journal.pone.0050876> (2012).
- Nishimura, K. *et al.* Identification of enzyme responsible for erythritol utilization and reaction product in yeast *Lipomyces starkeyi*. *J Biosci Bioeng* **101**, 303–308 (2006).
- Wojtatowicz, M. & Rymowicz, W. Effect of inoculum on kinetics and yield of citric acid production on glucose media by *Yarrowia lipolytica* A-101. *Acta Alimentaria Polonica (Poland)* **41**(2), 137–43 (1991).
- Rywińska, A., Rymowicz, W., Żarowska, B. & Wojtatowicz, M. Biosynthesis of Citric Acid from Glycerol by Acetate Mutants of *Yarrowia lipolytica* in Fed-Batch Fermentation. *Food Technology & Biotechnology* **47**, 1–6 (2009).
- Rywińska, A., Rymowicz, W., Żarowska, B. & Skrzypinski, A. Comparison of citric acid production from glycerol and glucose by different strains of *Yarrowia lipolytica*. *World Journal of Microbiology and Biotechnology* **26**, 1217–1224 (2010).
- Mirończuk, A. M., Dobrowolski, A., Rakicka, M., Rywińska, A. & Rymowicz, W. Newly isolated mutant of *Yarrowia lipolytica* MK1 as a proper host for efficient erythritol biosynthesis from glycerol. *Process Biochemistry* **50**, 61–68 (2015).
- Marchler-Bauer, A. *et al.* CDD: NCBI's conserved domain database. *Nucleic acids research* **43**, D222–226, <https://doi.org/10.1093/nar/gku1221> (2015).
- Marchler-Bauer, A. *et al.* CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic acids research* **45**, D200–d203, <https://doi.org/10.1093/nar/gkw1129> (2017).
- Gutierrez-Escribano, P. *et al.* The NDR/LATS kinase Cbk1 controls the activity of the transcriptional regulator Bcr1 during biofilm formation in *Candida albicans*. *PLoS pathogens* **8**, e1002683, <https://doi.org/10.1371/journal.ppat.1002683> (2012).
- Mirończuk, A. M., Rzechonek, D. A., Biegalska, A., Rakicka, M. & Dobrowolski, A. A novel strain of *Yarrowia lipolytica* as a platform for value-added product synthesis from glycerol. *Biotechnol Biofuels* **9**, 180, <https://doi.org/10.1186/s13068-016-0593-z> (2016).
- Sambrook, J. & Russell, D. W. *Molecular Cloning: A Laboratory Manual*. (Cold Spring Harbor Laboratory Press, (2001).
- Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)* **30**, 2114–2120, <https://doi.org/10.1093/bioinformatics/btu170> (2014).
- Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *Bioinformatics in Action* **17**, 10–12, doi:citeulike-article-id:11851772 (2012).
- Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)* **25**, 1754–1760 (2009).
- Devillers, H. *et al.* Draft Genome Sequence of *Yarrowia lipolytica* Strain A-101 Isolated from Polluted Soil in Poland. *Genome Announc* **4**, doi:<https://doi.org/10.1128/genomeA.01094-16> (2016).
- Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* **25**, 2078–2079 (2009).
- Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative C(T) method. *Nature protocols* **3**, 1101–1108 (2008).
- Nicaud, J.-M., Fabre, E. & Gaillardin, C. Expression of invertase activity in *Yarrowia lipolytica* and its use as a selective marker. *Current Genetics* **16**, 253–260 (1989).
- Fickers, P., Le Dall, M. T., Gaillardin, C., Thonart, P. & Nicaud, J. M. New disruption cassettes for rapid gene disruption and marker rescue in the yeast *Yarrowia lipolytica*. *J Microbiol Methods* **55**, 727–737 (2003).
- Mirończuk, A. M., Rakicka, M., Biegalska, A., Rymowicz, W. & Dobrowolski, A. A two-stage fermentation process of erythritol production by yeast *Y. lipolytica* from molasses and glycerol. *Bioresour Technol* **198**, 445–455 (2015).
- Mirończuk, A. M., Biegalska, A. & Dobrowolski, A. Functional overexpression of genes involved in erythritol synthesis in the yeast *Yarrowia lipolytica*. *Biotechnol Biofuels* **10**, 77 (2017).

Acknowledgements

We would like to thank Anna Biegalska for technical support in laboratory experiments. This study was financed by the Polish National Centre for Research and Development under project LIDER/010/207/L-5/13/NCBR/2014. Publication supported by Wrocław Centre of Biotechnology, programme the Leading Notional Research Centre (KNOW) for years 2014–2018.

Author Contributions

D.A.R. performed *in silico* analysis, cloning overexpression vector, shake-flask experiments and wrote the manuscript, H.D. performed gene mapping, C.N. performed gene mapping and revised the manuscript. W.R. revised the manuscript. A.M.M. designed the study, constructed the deletion cassette, performed R.T.-P.C.R. analysis and revised the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-12715-7>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017