

# SCIENTIFIC REPORTS



OPEN

## Discovery of a rhamnose utilization pathway and rhamnose-inducible promoters in *Pichia pastoris*

Bo Liu\*, Yuwei Zhang\*, Xue Zhang, Chengliang Yan, Yuhong Zhang, Xinxin Xu &amp; Wei Zhang

Received: 17 December 2015

Accepted: 12 April 2016

Published: 03 June 2016

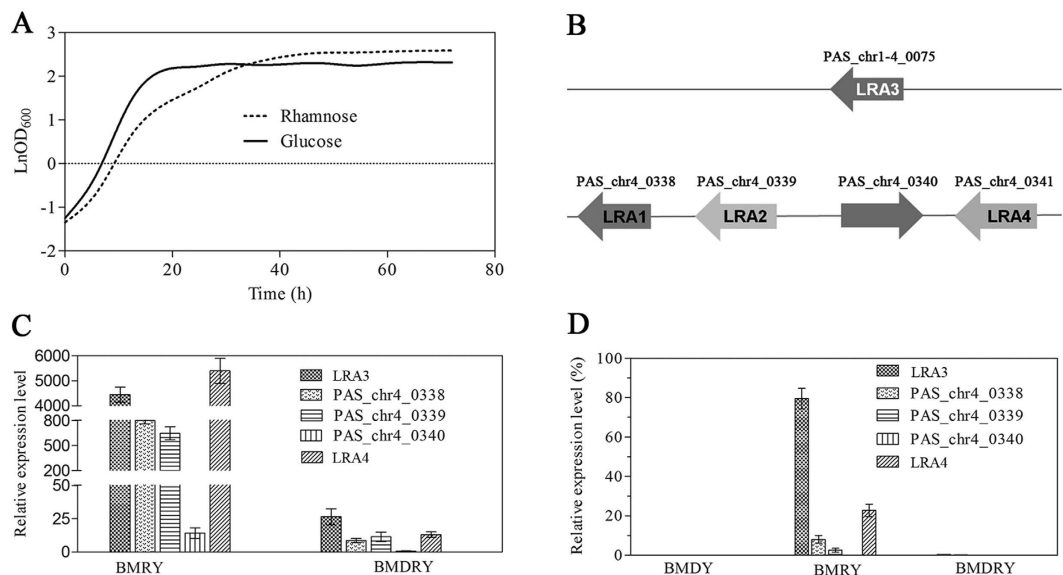
The rhamnose utilization pathway in *Pichia pastoris* has not been clarified although this strain can grow well on rhamnose as a sole carbon source. In this study, four genes, *PAS\_chr4\_0338*, *PAS\_chr4\_0339*, *PAS\_chr4\_0340*, and *PAS\_chr4\_0341*, were, for the first time, predicted to be involved in rhamnose metabolism along with the previously identified gene *PAS\_chr1\_4-0075*. Moreover, expression of these genes, especially *PAS\_chr4\_0341* and *PAS\_chr1\_4-0075* designated as *LRA4* and *LRA3*, was confirmed to significantly increase and clearly decrease in the presences of rhamnose and glucose, respectively. *LRA4* encoding a putative L-2-keto-3-deoxyrhamnonate aldolase, was further confirmed via gene disruption and gene complementation to participate in rhamnose metabolism. Using  $\beta$ -galactosidase and green fluorescent protein as reporters, the promoters of *LRA4* and *LRA3* performed well in driving efficient production of heterologous proteins. By using food grade rhamnose instead of the toxic compound methanol as the inducer, the two promoters would be excellent candidates for driving the production of food-grade and therapeutically important recombinant proteins.

L-rhamnose can be utilized for growth by multitudinous microorganisms as a sole carbon and energy source<sup>1–8</sup>, and to date two different L-rhamnose metabolism pathways have been clarified in prokaryotes and eukaryotes<sup>9,10</sup>. It was further found that expression of some genes related to L-rhamnose metabolism in some microorganisms, such as *Escherichia coli*, *Listeria monocytogenes*, and *Scheffersomyces (Pichia) stipitis*, was strictly induced by L-rhamnose and strongly repressed by D-glucose<sup>1,10,11,12</sup>. Hence, an abundance of rhamnose inducible promoters were exploited and adopted to drive the expression of target genes, for example, the *E. coli* rhaP<sub>BAD</sub> promoter performed excellently in the production of heterologous proteins in the presence of rhamnose<sup>4,13,14</sup>, and the commercial Expresso Rhamnose Cloning and Expression System was developed by the Lucigen Corporation. Therefore, rhamnose inducible promoters offer great potential for the production of target proteins and many more promoters of this kind should be exploited to satisfy the demand in basic research and for practical applications.

*Pichia pastoris*, a methylotrophic yeast, is an outstanding host for heterologous protein production. Several strong methanol-inducible promoters of genes related to the methanol utilization pathway, especially the well-characterized alcohol oxidase 1 promoter (*P<sub>AOX1</sub>*), have been frequently used to express recombinant proteins<sup>15–17</sup>. However, this promoter is not perfect for recombinant protein production due to two disadvantages of the inducer, methanol: it is highly flammable and hazardous to health. Thus, other inducible strong promoters are in high demand for controllable expression of heterologous genes in *P. pastoris*.

*P. pastoris* can grow on rhamnose as a sole carbon source, but the genes associated with L-rhamnose metabolism have not been identified. We hypothesized that the promoters of L-rhamnose metabolism-related genes would also perform well in realizing massive production of heterologous proteins. This study aimed to disclose the putative genes involved in rhamnose metabolism pathways in *P. pastoris* and to exploit the promoters of these genes for controllable production of heterologous proteins. As expected, a series of putative genes involved in L-rhamnose metabolism in *P. pastoris* was preliminarily identified and the promoters isolated from these genes, especially *PAS\_chr4\_0341* and *PAS\_chr1-4\_0075* designated as *LRA4* and *LRA3*, exhibited obvious inducibility by rhamnose. The two promoters could be exploited to enable efficient production of target proteins, especially food-grade and therapeutically important recombinant proteins, in *P. pastoris*.

Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, China. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to W.Z. (email: zhangwei02@caas.cn)



**Figure 1. Characterization of putative genes involved in L-rhamnose metabolism in *P. pastoris*.**

(A) Growth profiles of *P. pastoris* grown on 2% (w/v) L-rhamnose or glucose as a sole carbon and energy source. (B) Relative positions of putative genes involved in L-rhamnose metabolism in the chromosome of *P. pastoris*. LRA1, L-rhamnose 1-dehydrogenase; LRA2, L-rhamnono- $\gamma$ -lactonase; LRA3, L-rhamnonate dehydratase; LRA4, L-KDR aldolase. (C,D) Relative transcription profiles of these genes in *P. pastoris* grown in different media. The relative expression value of each gene in BMDY was assigned as 1 and the GAPDH gene was used as a reference (C); the relative expression value of the GAPDH gene was used as a reference and was assigned as 100% (D). Each test was conducted in triplicate and the results are presented as means  $\pm$  SEM of three replicates ( $n = 3$ ).

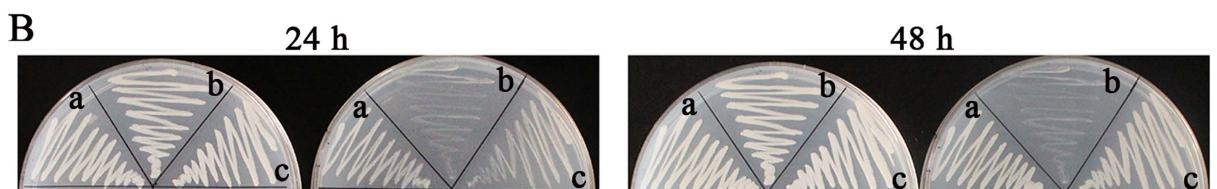
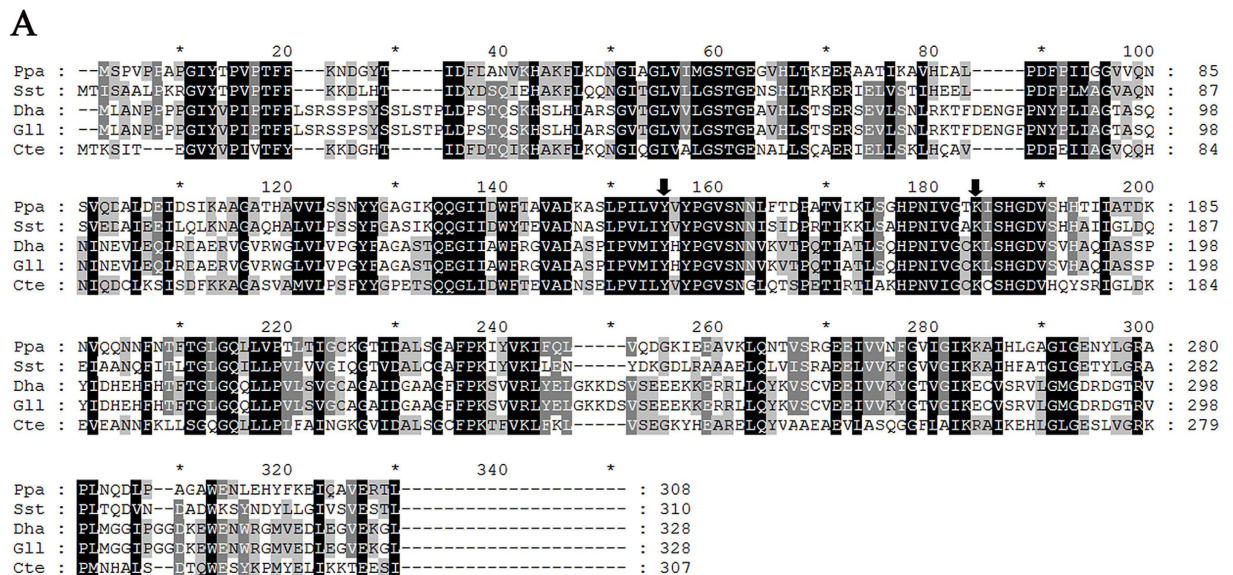
## Results

**Putative genes involved in L-rhamnose metabolism in *P. pastoris*.** *P. pastoris* grew slightly slower on 2% (w/v) L-rhamnose than on glucose (Fig. 1A), suggesting that *P. pastoris* can utilize L-rhamnose as a sole carbon source. However, to date, genes involved in L-rhamnose metabolism have not been elucidated in *P. pastoris* except for LRA3, which was annotated as L-rhamnonate dehydratase after genome sequencing. By searching the nucleotide sequences of genes homologous to those of genes related to rhamnose metabolism in *S. stipitis*, the four following genes were predicted to participate in rhamnose utilization in *P. pastoris*: PAS\_chr4\_0338, PAS\_chr4\_0339, PAS\_chr4\_0340, and LRA4, which putatively encoded L-rhamnose 1-dehydrogenase, L-rhamnono- $\gamma$ -lactonase, a transcription regulator, and L-2-keto-3-deoxyrhamnonate (L-KDR) aldolase, respectively (Fig. 1B).

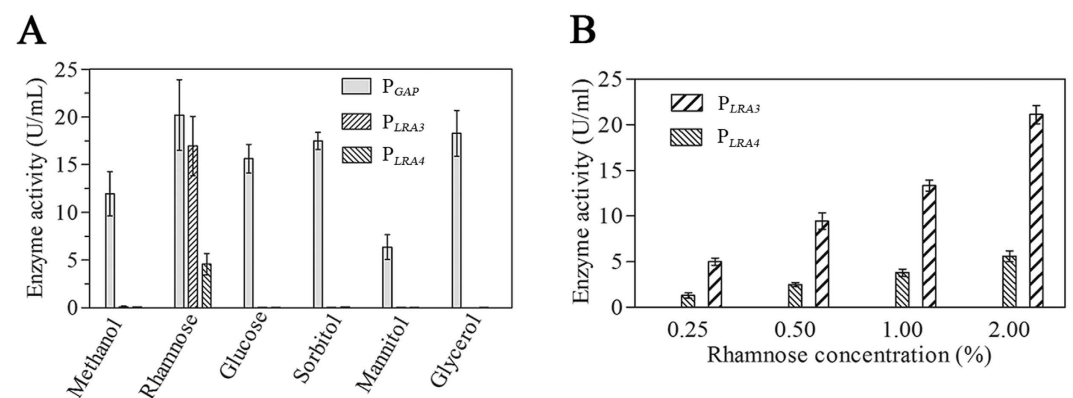
To further investigate the relationships between these genes and rhamnose metabolism, *P. pastoris* was cultured in BMDY, BMR, and BMDY media. When rhamnose was used as a sole carbon source, transcription levels of all four genes and LRA3 increased. In particular, transcription levels of LRA3 and LRA4 were enhanced over 4000-fold compared to *P. pastoris* growing on glucose. When a mixture of rhamnose and glucose was used as the carbon source, the expression levels of these genes clearly decreased (Fig. 1C). The above results showed that expression of the five genes was induced by rhamnose and repressed by glucose, and also indicated that these genes are probably related to rhamnose metabolism. Moreover, transcriptional levels of LRA3 and LRA4 were respectively 75% and 22% of glyceraldehyde 3-phosphate dehydrogenase (GAPDH) transcription levels in the presence of an initial concentration of 2% rhamnose (Fig. 1D). Additionally, the above results also imply that the LRA3 promoter ( $P_{LRA3}$ ) was strongly rhamnose-inducible, while the LRA4 promoter ( $P_{LRA4}$ ) was weakly rhamnose-inducible.

LRA4 is a homolog of L-KDR aldolase from *S. stipitis* according to amino acid sequence identity. In addition, LRA4 showed high homology when aligned with aldolases from *Debaryomyces hansenii*, *Candida tenuis* and *Glarea lozoyensis*. Furthermore, two amino acid residues (a lysine residue and a tyrosine residue) that are essential active sites in L-KDR aldolase from *S. stipitis*<sup>10</sup> are also present in LRA4 (Fig. 2A). In order to confirm that LRA4 was associated with L-rhamnose metabolism, a LRA4-disrupted strain and a LRA4-complemented strain were developed by gene disruption and gene complementation, respectively. The wild-type strain, LRA4-disrupted strain, and LRA4-complemented strain were grown on glucose or rhamnose as a sole carbon source. As expected, all three strains grew well on glucose, but only the wild-type strain and the LRA4-complemented strain grew well on rhamnose (Fig. 2B), which suggests that LRA4 is definitely involved in L-rhamnose utilization.

**Expression of heterologous genes driven by  $P_{LRA3}$  and  $P_{LRA4}$  in *P. pastoris*.** Previously, LRA4 and LRA3 exhibited relatively high transcription activities in the presence of rhamnose. To investigate whether  $P_{LRA3}$  and  $P_{LRA4}$  were subject to strict rhamnose regulation, the effects of several carbon sources on expression of *lacB* under the control of  $P_{LRA3}$  and  $P_{LRA4}$  were monitored. Expression of *lacB* driven by  $P_{LRA3}$  and  $P_{LRA4}$  was only observed when using rhamnose as a sole carbon source. In contrast, under control of the well-characterized



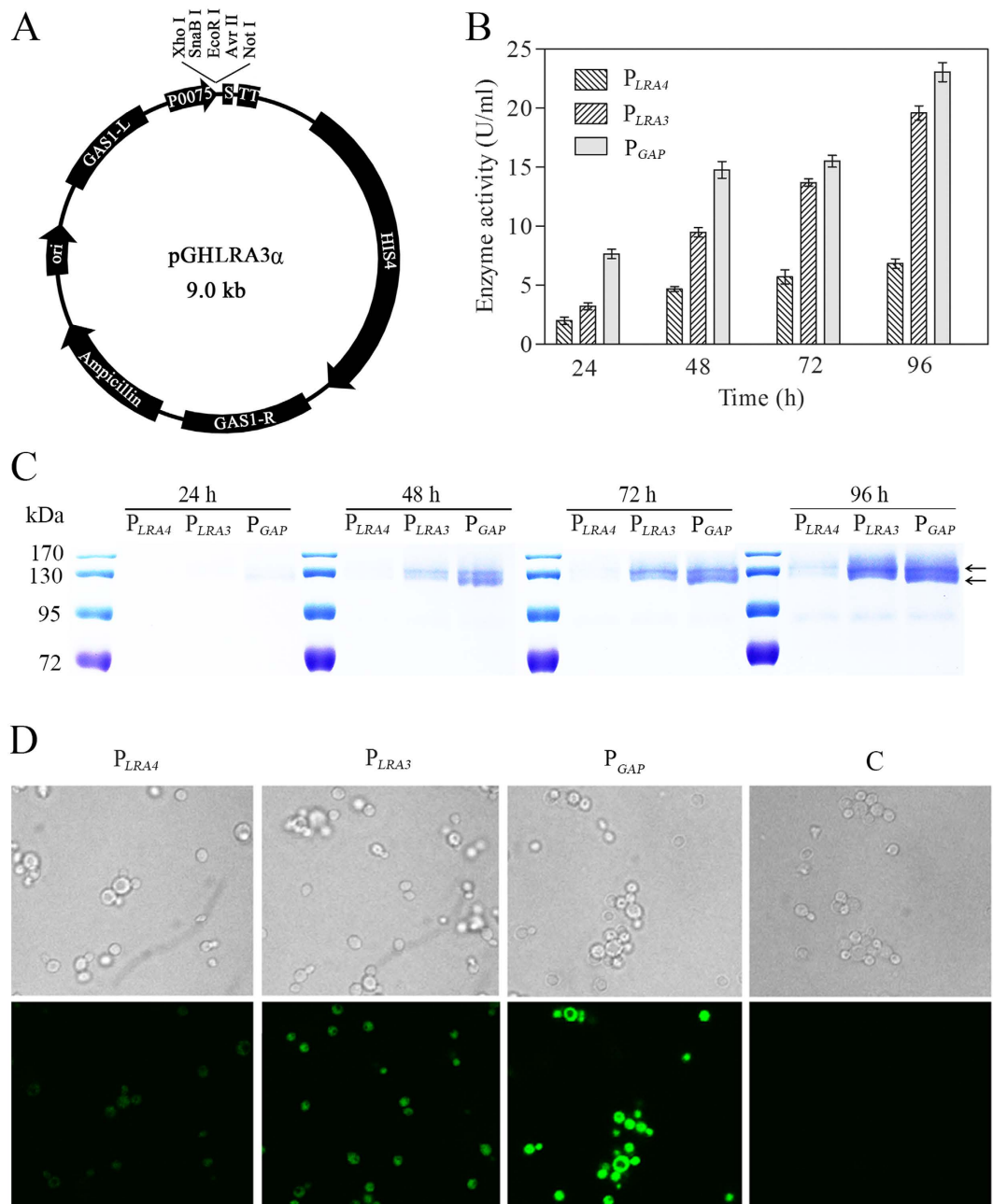
**Figure 2. Physiological function assay of *LRA4*.** (A) Multiple amino acid sequence alignment of L-KDR aldolases from different strains. *Ppa*, *Pichia pastoris*; *Sst*, *Scheffersomyces stipitis*; *Dha*, *Debaryomyces hansenii*; *Gll*, *Glarea lozoyensis*; *Cte*, *Candida tenuis*. The two catalytic residues (a lysine residue and a tyrosine residue) in L-KDR aldolase are indicated with arrows. (B) Growth profiles of the wild-type strain, *LRA4*-disrupted strain, and *LRA4*-complemented strain grown on glucose (left) or rhamnose (right) as a sole carbon source. (a) Wild-type strain, (b) *LRA4*-disrupted strain, (c) *LRA4*-complemented strain.



**Figure 3. Expression profiles of  $\beta$ -galactosidase driven by  $P_{LRA3}$  and  $P_{LRA4}$  in the presence of different carbon sources (A) and various rhamnose concentrations (B).**

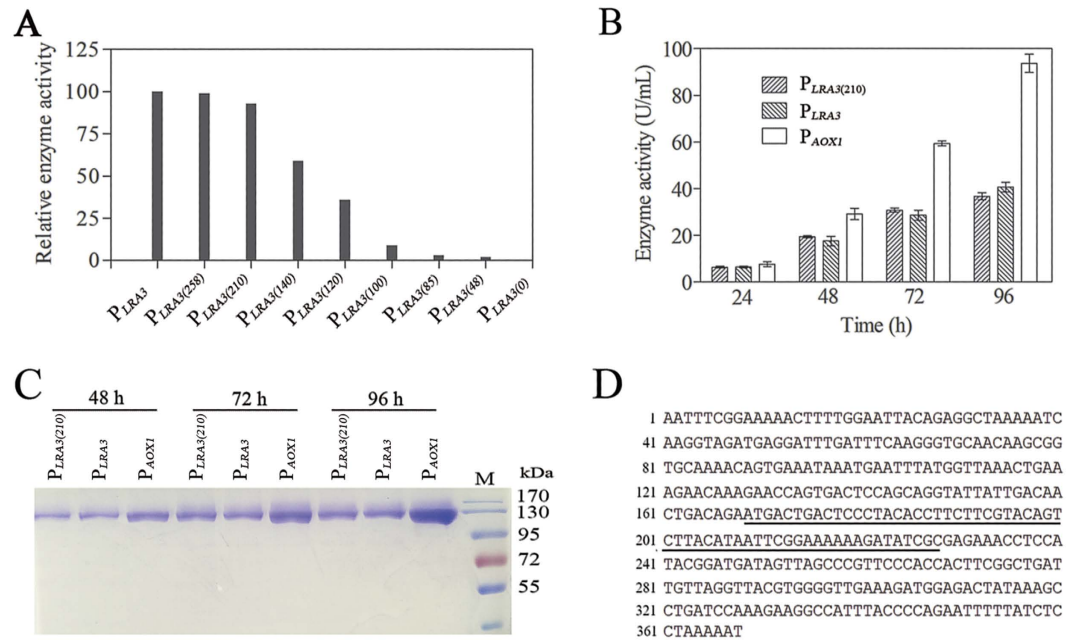
*GADPH* promoter ( $P_{GAP}$ ) *lacB* expression was observed when methanol, glucose, mannitol, sorbitol, and glycerol were used as sole carbon sources, in addition to rhamnose (Fig. 3A). The above results indicate that  $P_{LRA3}$  and  $P_{LRA4}$  were strict rhamnose-dependent promoters. Additionally, a dose-effect relationship occurred between rhamnose concentration (0.25–2.0%) and production of  $\beta$ -galactosidase driven by  $P_{LRA3}$  and  $P_{LRA4}$  (Fig. 3B), meaning that production of a target protein could be controlled in a rhamnose-dose dependent manner.

To validate the efficiency of the rhamnose-inducible  $P_{LRA3}$  and  $P_{LRA4}$  in target gene expression,  $P_{GAP}$  and the two identified rhamnose-inducible promoters were subjected to a comparative protein production study using  $\beta$ -galactosidase and green fluorescent protein (GFP) as reporters. Vectors pGHLRA3 $\alpha$ , pGHLRA4 $\alpha$ , and pGH-GAP $\alpha$  based on  $P_{LRA3}$ ,  $P_{LRA4}$ , and  $P_{GAP}$ , respectively, were developed, with pGHLRA3 $\alpha$  illustrated as an example



**Figure 4.** Expression profiles of heterologous genes driven by  $P_{LRA3}$  and  $P_{LRA4}$  in *P. pastoris*. (A) Schematic representation of plasmid pGHLRA3 $\alpha$  for heterogeneous gene expression in *P. pastoris*. (B)  $\beta$ -galactosidase activity under the control of  $P_{LRA3}$ ,  $P_{LRA4}$ , and  $P_{GAP}$  at various time points. (C) Production of recombinant  $\beta$ -galactosidase using  $P_{LRA3}$ ,  $P_{LRA4}$ , and  $P_{GAP}$  for various durations. (D) Green fluorescence under the control of  $P_{LRA3}$ ,  $P_{LRA4}$ , and  $P_{GAP}$ ; 'C' represents *P. pastoris* GS115, which was used as a control.

in Fig. 4A. One reporter gene, *lacB*, encoding a  $\beta$ -galactosidase, was ligated into the above vectors via the *SnaBI* and *NotI* restriction sites to generate pGHLRA3 $\alpha$ LacB, pGHLRA4 $\alpha$ LacB, and pGHGAP $\alpha$ LacB, in which *lacB* expression was controlled by  $P_{LRA3}$ ,  $P_{LRA4}$ , and  $P_{GAP}$ , respectively, and was then integrated into the *gasI* locus of the *P. pastoris* chromosome after transformation. The  $\beta$ -galactosidase activities in culture supernatants of the three recombinant *P. pastoris* strains grown in BMRY and BMDY were monitored. It was observed that  $\beta$ -galactosidase activities in supernatants of the different cultures in BMRY increased with incubation duration and reached a maximum at 72 h: 23 U/ml, 18 U/ml, and 6 U/ml for recombinant strains harboring pGHLRA3 $\alpha$ LacB, pGHLRA4 $\alpha$ LacB, and pGHGAP $\alpha$ LacB, respectively (Fig. 4B). The target protein productions under the control of  $P_{LRA3}$  and  $P_{LRA4}$  in the presence of rhamnose were approximately 80% and 26% of that under  $P_{GAP}$ , respectively, which was consistent with the transcription levels determined by real-time PCR. Production of  $\beta$ -galactosidase in culture supernatants of the different recombinant strains was assayed via sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Fig. 4C).



**Figure 5. Identification of the *LRA3* minimal promoter.** (A) Promoter activities of all 5'-deletion promoter constructs were evaluated according to  $\beta$ -galactosidase activity in culture supernatant of recombinant strains grown in BMRY. Numbers refer to the number of nucleotides remaining in the 5'-flanking region upstream of the ATG of the initiating methionine of *LRA3*. (B)  $\beta$ -galactosidase activity and (C)  $\beta$ -galactosidase production in culture supernatant of recombinant strains harboring the indicated promoter to drive *lacB* expression at different induction time. (D) The nucleotide sequence of putative P<sub>LRA3</sub> and the region between -140 to -210 bp (underlined). Each test was conducted in triplicate and the results are presented as means.

To investigate whether the two promoters also drove efficient expression of other genes, expression of another reporter gene, GFP (*gfp*), was also monitored. For intracellular expression of *gfp*, three new vectors, pGHLRA3, pGHLRA4, and pGHGAP, were developed without a protein secretion signal based on vectors pGHLRA3 $\alpha$ , pGHLRA4 $\alpha$ , and pGHGAP $\alpha$ . *gfp* was ligated into vectors pGHLRA3, pGHLRA4, and pGHGAP via the *EcoRI* and *NotI* restriction sites, and was integrated into the *gas1* locus of the *P. pastoris* chromosome. Green fluorescence was monitored after the recombinant strains were grown in BMRY. Fluorescence was observed in all three recombinant strains, in which *gfp* expression was regulated by P<sub>GAP</sub>, P<sub>LRA3</sub>, and P<sub>LRA4</sub>. Mean green fluorescence intensity was 235 (P<sub>LRA4</sub>), 563 (P<sub>LRA3</sub>) and 2249 (P<sub>GAP</sub>) respectively, which indicated that the strengths of the three promoters were ranked as follows: P<sub>GAP</sub> (high), P<sub>LRA3</sub> (middle), and P<sub>LRA4</sub> (weak) (Fig. 4D). According to the above results, P<sub>LRA3</sub> would be more advantageous for inducible expression of target genes than P<sub>LRA4</sub>.

**Characterization of putative P<sub>LRA3</sub>.** P<sub>LRA3</sub>, a newly discovered strong rhamnose-inducible promoter, could potentially be widely used to drive production of proteins of interest in *P. pastoris*. To identify the minimal promoter region, 5'-deletion constructs of P<sub>LRA3</sub> were derived from pGHLRA3 $\alpha$ LacB and were inserted into pGHLRA3 $\alpha$ LacB between the *AsiI* and *SnaBI* sites, generating pGHLRA3(258) $\alpha$ LacB, pGHLRA3(210) $\alpha$ LacB, pGHLRA3(140) $\alpha$ LacB, pGHLRA3(120) $\alpha$ LacB, pGHLRA3(100) $\alpha$ LacB, pGHLRA3(85) $\alpha$ LacB, pGHLRA3(48) $\alpha$ LacB, and pGHLRA3(0) $\alpha$ LacB. These deletion constructs were transformed into *P. pastoris* and the  $\beta$ -galactosidase activities in culture supernatants of the different recombinant strains were determined. The results showed that a promoter harboring a 5' flanking region greater than 210 bp upstream of the ATG of the initiating methionine of *LRA3* was sufficient for transcriptional activation (Fig. 5A), which suggests that a 210-bp DNA fragment harbors all the necessary elements to express recombinant proteins in *P. pastoris*. The promoter with a 5' flanking region containing the 140 bp upstream of the translation start site was not as efficient, indicating that elements in the DNA region between -140 bp and -210 bp were essential to efficient expression of target genes. Additionally, some important regulatory elements regulating gene expression occur between -100 bp and -140 bp, as the deletion of this region resulted in an obvious decrease in  $\beta$ -galactosidase activity (Fig. 5A).

The efficiency of P<sub>LRA3(210)</sub> and P<sub>LRA3</sub> was further determined using  $\beta$ -galactosidase as a reporter gene and P<sub>AOX1</sub> as a positive control promoter. The  $\beta$ -galactosidase activities in culture supernatants of recombinant strains harboring pGHLRA3 $\alpha$ LacB, pGHLRA3(210) $\alpha$ LacB and pPIC9LacB increased during the induction. At 96 h, activity and production of  $\beta$ -galactosidase in culture supernatants of strain harboring pGHLRA3 $\alpha$ LacB and pGHLRA3(210) $\alpha$ LacB were almost same, and were less than half of that of strain harboring pPIC9LacB (Fig. 5B,C). The results further indicated that P<sub>LRA3</sub> was a middle strong rhamnose-inducible promoter. The nucleotide sequence of P<sub>LRA3</sub> was shown and the region between -140 to -210 bp was underlined (Fig. 5D).

## Discussion

Identification of genes related to rhamnose metabolism in *P. pastoris* allowed us to further elucidate the rhamnose metabolism pathway and to exploit inducible promoters for controllable expression of target genes in *P. pastoris*. In this study, a set of genes involved in rhamnose metabolism in *P. pastoris* was first predicted according to the nucleotide sequence identity to genes in *S. stipitis*. Then, *LRA4* was further confirmed to correlate with rhamnose metabolism based on the result that *LRA4*-disrupted *P. pastoris* did not grow on rhamnose as a sole carbon source, while the wild-type strain and the *LRA4*-complemented strain did. To further confirm the physiological role of this gene, the biochemical characterization of the enzyme encoded by *LRA4* was examined in another study. Similarly, the physiological roles of the other genes we identified should also be verified via gene disruption and complementation and biochemical function assays in future studies. Additionally, it was noted that the genes encoding the enzymes for rhamnose metabolism in *P. pastoris* did not cluster. Three genes (*PAS\_chr4\_0338*, *PAS\_chr4\_0339*, and *LRA4*) and the regulator gene (*PAS\_chr4\_0340*) were organized in a gene cluster, while *LRA3* was independent. This is not the same organization found in *S. stipitis*, in which a set of genes correlated with rhamnose catabolism clustered in a rhamnose operon. The difference reflects different evolutionary processes in rhamnose metabolism between *S. stipitis* and *P. pastoris*.

$P_{AOXI}$ , which is strictly and strongly activated by methanol, was the promoter most widely used for production of heterologous proteins in *P. pastoris*. However, strong promoters that produce heterologous proteins in high quantities are unsuitable for certain situations, such as production of active products that are highly dependent on correct folding or processing in the secretory pathway<sup>18</sup> or the expression of various lethal or toxic genes; in these cases weaker promoters have been reported to be more feasible<sup>19,20</sup>. Also, in addition to using the toxic compound methanol as the inducer,  $P_{AOXI}$ -based fermentation includes the following disadvantages: 1) relatively long fermentation duration and 2) sophisticated feeding strategies including but not limited to, *i*) initial culture of recombinants in buffered glycerol medium, followed by *ii*) harvesting and resuspending cells in methanol medium. To broaden the applications of *P. pastoris* and simplify the process of fermentation, alternative promoters that do not require methanol are in demand, and various promoters have been described<sup>21</sup>. Five rhamnose-inducible genes were discovered in our study, and among them  $P_{LRA3}$  and  $P_{LRA4}$  were strictly activated in the presence of rhamnose and almost completely repressed by glucose.  $P_{LRA3}$  in the presence of rhamnose was confirmed to be comparable to the typical strong promoter,  $P_{GAP}$ , which has been commonly used to drive efficient expression of target genes in *P. pastoris*. Due to using rhamnose instead of the flammable and toxic compound methanol as an inducer,  $P_{LRA3}$  could be used for substantial production of recombinant proteins in the food industry and production of therapeutically important recombinant proteins for commercial applications. As for  $P_{LRA4}$ , it was much weaker than  $P_{LRA3}$  and is a better candidate for certain academic purposes, for example, controllably expressing potentially lethal or toxic genes that affect cell viability. Few promoters isolated from genes related to sugar metabolism have been reported in *P. pastoris* and, to our knowledge, the sorbitol dehydrogenase promoter, a constitutively expressed promoter with a transcription strength similar to  $P_{GAP}$ , is the only one that had been intensively investigated prior to our study. Compared to the sorbitol dehydrogenase promoter,  $P_{LRA3}$  is disadvantageous in its transcription activity, however  $P_{LRA3}$  is advantageous for controllable expression of heterologous genes, including some genes in which the products might be toxic to the host cell.

## Materials and Methods

**Bacterial strains and media.** *E. coli* Trans1-T1 (TransGen, Beijing, China) and *P. pastoris* GS115 (Invitrogen, Carlsbad, CA, USA) were used as the hosts for gene cloning and expression, respectively. The *pEasy-Blunt* Simple Cloning Vector was purchased from TransGen.

MD medium contained 1.34% yeast nitrogen base (YNB),  $4 \times 10^{-5}$ % biotin, 2% dextrose, and 2% agar; BMDY, BMRY, and BMDRY media (pH 6.0) contained 300 mM potassium phosphate, 1.34% YNB,  $4 \times 10^{-5}$ % biotin, and  $4 \times 10^{-3}$ % histidine, plus either 2% dextrose (BMDY), 2% rhamnose (BMRY), or 1% dextrose and 1% rhamnose (BMDRY).

The primers used for PCR are listed in Supplementary Table S1.

**Total RNA preparation and real-time PCR.** *P. pastoris* GS115 was grown in BMDY, BMRY, or BMDRY to OD<sub>600</sub> 1.2–1.4, and then cells were harvested by centrifugation ( $12,000 \text{ g} \times 5 \text{ min}$ ) at 4 °C and immediately stored at –80 °C. Total RNA from *P. pastoris* cells was extracted with Trizol reagent (Invitrogen) followed by digestion with DNase I (RNase-free; MBI Fermentas, St. Leon-Roth, Germany) to remove trace DNA. cDNA was synthesized from each RNA sample (1 μg) using the RevertAid H Minus First Strand cDNA Synthesis Kit (K1631; MBI Fermentas). Real-time PCR assays were performed in the IQ5 Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, CA) using the SuperReal PreMix SYBR Green I Kit (Toyobo, Osaka, Japan). Primers used for real-time PCR are listed in Table S1. The GAPDH gene was used as a reference. All samples were analyzed in triplicate and the data are presented as means ± standard deviation (SD).

**Construction of a homologous recombination vector in *P. pastoris*.** The genomic DNA was isolated from *P. pastoris* cells using the TIANamp Yeast DNA Kit (Tiangen, Beijing, China). Based on the genome sequence of *P. pastoris* GS115, DNA fragments upstream and downstream of the open reading frame of *gas1* were amplified from genomic DNA by two pairs of primers (*gas1*-F1 and *gas1*-R1, *gas1*-F2 and *gas1*-R2, respectively). *HIS4*, a selectable marker for isolating *Pichia* recombinant strains, was amplified from plasmid pPIC9 using the primers *his4*-F and *his4*-R. The three abovementioned DNA fragments were ligated into a DNA fragment named “GH” via overlap-extension PCR with the primers *gas1*-LF and *gas1*-RR. In addition, a DNA fragment designated “OA” containing the origin of replication derived from pBR322 and an ampicillin resistance gene was amplified from plasmid pPIC9 with primers OA-F and OA-R. Finally, the two DNA fragments, GH and OA, were ligated

into a circular molecule by homologous recombination using the CloneEZ Kit (GenScript, Nanjing, China), generating the recombination vector pGH01. pGH01 was integrated into the *gas1* locus of the *P. pastoris* chromosome via homologous recombination.

**LRA4 disruption and complementation.** DNA fragments upstream and downstream of *LRA4*, which were used as homologous arms for integration of target genes into the *LRA4* locus, were amplified from genomic DNA of *P. pastoris* GS115. The expression cassette of a zeocin resistance gene was obtained from the plasmid pPICz $\alpha$ . The three PCR fragments were fused by overlap-extension PCR to generate a gene disruption cassette for disruption of *LRA4* in *P. pastoris*. From the resultant DNA fragments, 500 ng were transformed into *P. pastoris* GS115 via electroporation. Transformants were first screened on yeast extract-peptone-dextrose (YPD) agar plates containing 100  $\mu$ g/ml zeocin and then verified by PCR.

To complement *LRA4*, the complementation plasmid pGH/*LRA4* was developed as follows. A DNA fragment harboring full length *LRA4* was cloned from genomic DNA of *P. pastoris* GS115 and then ligated into the pGH01 vector via the *AscI* and *PacI* restriction sites to generate the pGH/*LRA4* plasmid. pGH/*LRA4* was introduced into *P. pastoris* GS115 with *LRA4*-disruption via electroporation after linearization by the restriction enzyme *SwaI*. Transformants were screened on MD medium and then verified by PCR and DNA sequencing.

**Expression of exogenous genes under  $P_{LRA3}$  and  $P_{LRA4}$  in *P. pastoris*.** DNA fragments containing the putative promoter regions of *LRA3*, *LRA4*, and *GAPDH* were obtained from the genomic DNA of *P. pastoris*. The DNA fragment comprising the protein secretion signal, multiple cloning site, and transcription termination was cloned from plasmid pPIC9. The two DNA fragments were fused by overlap-PCR to generate the exogenous gene expression cassette in which exogenous gene expression was driven by  $P_{LRA3}$ ,  $P_{LRA4}$  and  $P_{GAP}$ , and then the cassettes were respectively ligated into vector pGH01 via the *AscI* and *PacI* restriction sites to generate expression vectors pGHLRA3 $\alpha$ , pGHLRA4 $\alpha$ , and pGHGAP $\alpha$ . Additionally, three vectors, pGHLRA3, pGHLRA4, and pGHGAP, were developed without the protein secretion signal in the same manner as pGHLRA3 $\alpha$ , pGHLRA4 $\alpha$ , and pGHGAP $\alpha$ .

The *lacB* gene (GenBank accession number: AJ431643.1), which was cloned from the *Aspergillus candidus* chromosome and encodes a  $\beta$ -galactosidase, was ligated into pGHLRA3 $\alpha$ , pGHLRA4 $\alpha$ , and pGHGAP $\alpha$  and pPIC9 via the *SnaBI* and *NotI* restriction sites. The *gfp* gene (GenBank accession number: JQ733033.1) encoding green fluorescent protein was synthesized by GenScript Biotechnology Co. Ltd. and was ligated into pGHLRA3, pGHLRA4, and pGHGAP via the *EcoRI* and *NotI* restriction sites.

**Generation of 5'-serially deleted  $P_{LRA3}$ .** The 5'-serially deleted  $P_{LRA3}$  constructs were generated by PCR using pGHLRA3 $\alpha$ LacB as a template. The DNA fragment harboring  $P_{LRA3}$  and partial *lacB* in pGHLRA3 $\alpha$ LacB was replaced by the 5'-serially deleted promoters and partial *lacB* between restriction sites *AscI* and *SnaBI* to generate plasmids pGHLRA3(256) $\alpha$ LacB, pGHLRA3(210) $\alpha$ LacB, pGHLRA3(140) $\alpha$ LacB, pGHLRA3(120) $\alpha$ LacB, pGHLRA3(100) $\alpha$ LacB, pGHLRA3(85) $\alpha$ LacB, pGHLRA3(48) $\alpha$ LacB, and pGHLRA3(0) $\alpha$ LacB.

***lacB* expression and enzyme assays.** To evaluate the strength of  $P_{LRA3}$ ,  $P_{LRA4}$  and  $P_{GAP}$ , a single colony of different recombinant *Pichia* cells was inoculated into in YPD medium and grown overnight at 28 °C with vigorous shaking (200 rpm). The cultures were inoculated into fresh BMRY medium at 1% (v/v) and grown at 28 °C with vigorous shaking (200 rpm).

To evaluate the strength of  $P_{LRA3}$ ,  $P_{LRA3(210)}$  and  $P_{AOX1}$ , a single colony of different recombinant *Pichia* cells was inoculated into in 50-ml YPD medium at 28 °C for 48 h in shaking flasks. The cells with *lacB* expression under the control of  $P_{LRA3}$  and  $P_{LRA3(210)}$  were harvested and then grown in 50-ml BMRY medium, and the cells with *lacB* expression under the control of  $P_{AOX1}$  were harvested and then grown in 50-ml BMMY medium with addition of methanol to a final concentration of 1.0% [v/v] at an interval of 24 h.  $\beta$ -galactosidase activity and production in culture supernatant were monitored at intervals.

The  $\beta$ -galactosidase activity was assayed as described by Nie *et al.*<sup>15</sup>, with minor modification. The reaction system containing 200  $\mu$ l of enzyme solution and 800  $\mu$ l of 0.25% (w/v) ortho-nitrophenyl- $\beta$ -galactoside (oNPG) in pre-incubated (60 °C for 5 min) phosphate-citrate buffer (50 mM, pH 5.2) was incubated at 60 °C for 15 min, followed by sequential addition of 1 ml of 10% trichloroacetic acid and 2 ml of 1 M Na<sub>2</sub>CO<sub>3</sub>. The absorbance at 420 nm was measured. One unit of  $\beta$ -galactosidase was defined as the amount of enzyme that released 1  $\mu$ mol of *o*-nitrophenol per minute under standard conditions (pH 5.2, 60 °C, 15 min).

***gfp* expression and assays of green fluorescence intensity.** A single colony of recombinant *Pichia* cells was inoculated into in YPD medium and grown overnight at 28 °C with vigorous shaking (200 rpm). The cultures were inoculated into fresh BMRY medium at 1% (v/v) and grown for 48 h at 28 °C with vigorous shaking (200 rpm). Recombinant *Pichia* cells were then collected, washed three times, and resuspended in 0.9% NaCl. Green fluorescence intensity in recombinant cells was determined using a laser scanning confocal microscope (Nikon, Japan) with 2s per image to collect a 1024  $\times$  1024 pixel image.

## References

1. Koivistoinen, O. M. *et al.* Characterisation of the gene cluster for l-rhamnose catabolism in the yeast *Scheffersomyces (Pichia) stipitidis*. *Gene* **492**, 177–185 (2012).
2. Koivistoinen, O. M. *et al.* Identification in the yeast *Pichia stipitidis* of the first L-rhamnose-1-dehydrogenase gene. *FEBS J* **275**, 2482–2488 (2008).
3. Tobin, J. F. & Schleif, R. F. Purification and properties of RhaR, the positive regulator of the L-rhamnose operons of *Escherichia coli*. *J Mol Biol* **211**, 75–89 (1990).

4. Cardona, S. T., Mueller, C. L. & Valvano, M. A. Identification of essential operons with a rhamnose-inducible promoter in *Burkholderia cenocepacia*. *Appl Environ Microbiol* **72**, 2547–2555 (2006).
5. Petit, E. *et al.* Involvement of a bacterial microcompartment in the metabolism of fucose and rhamnose by *Clostridium phytofermentans*. *PLoS One* **8**, e54337 (2013).
6. Madduri, K., Waldron, C. & Merlo, D. J. Rhamnose biosynthesis pathway supplies precursors for primary and secondary metabolism in *Saccharopolyspora spinosa*. *J Bacteriol* **183**, 5632–5638 (2001).
7. Obradors, N., Badia, J., Baldoma, L. & Aguilar, J. Anaerobic metabolism of the L-rhamnose fermentation product 1,2-propanediol in *Salmonella typhimurium*. *J Bacteriol* **170**, 2159–2162 (1988).
8. Levinson, S. L. & Krulwich, T. A. Metabolism of L-rhamnose in *Arthrobacter pyridinolis*. *J Gen Microbiol* **96**, 277–286 (1976).
9. Watanabe, S. & Makino, K. Novel modified version of nonphosphorylated sugar metabolism—an alternative L-rhamnose pathway of *Sphingomonas* sp. *FEBS J* **276**, 1554–1567 (2009).
10. Watanabe, S., Saimura, M. & Makino, K. Eukaryotic and bacterial gene clusters related to an alternative pathway of nonphosphorylated L-rhamnose metabolism. *J Biol Chem* **283**, 20372–20382 (2008).
11. Haldimann, A., Daniels, L. L. & Wanner, B. L. Use of new methods for construction of tightly regulated arabinose and rhamnose promoter fusions in studies of the *Escherichia coli* phosphate regulon. *J Bacteriol* **180**, 1277–1286 (1998).
12. Fieseler, L., Schmitter, S., Teiserskas, J. & Loessner, M. J. Rhamnose-inducible gene expression in *Listeria monocytogenes*. *PLoS One* **7**, e43444 (2012).
13. Cardona, S. T. & Valvano, M. A. An expression vector containing a rhamnose-inducible promoter provides tightly regulated gene expression in *Burkholderia cenocepacia*. *Plasmid* **54**, 219–228 (2005).
14. Giacalone, M. J. *et al.* Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system. *BioTechniques* **40**, 355–364 (2006).
15. Nie, C. *et al.* Production and secretion of *Lactobacillus crispatus* beta-galactosidase in *Pichia pastoris*. *Protein Express Purif* **92**, 88–93 (2013).
16. Hao, J., Xu, L., He, H., Du, X. & Jia, L. High-level expression of *Staphylococcal* protein A in *Pichia pastoris* and purification and characterization of the recombinant protein. *Protein Express Purif* **90**, 178–185 (2013).
17. Lee, P. Y., Yong, V. C., Rosli, R., Gam, L. H. & Chong, P. P. Cloning, expression and purification of squalene synthase from *Candida tropicalis* in *Pichia pastoris*. *Protein Express Purif* **94**, 15–21 (2014).
18. Vogl, T. & Glieder, A. Regulation of *Pichia pastoris* promoters and its consequences for protein production. *New Biotechnol* **30**, 385–404 (2013).
19. Ruth, C. *et al.* Variable production windows for porcine trypsinogen employing synthetic inducible promoter variants in *Pichia pastoris*. *Syst Synth Biol* **4**, 181–191 (2010).
20. Cereghino, J. L. & Cregg, J. M. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol Rev* **24**, 45–66 (2000).
21. Ahmad, M., Hirz, M., Pichler, H. & Schwab, H. Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol* **98**, 5301–5317 (2014).

## Acknowledgements

This research was supported by the National High Technology Research and Development Program of China (2012AA022105, 2012AA022207).

## Author Contributions

B.L. and W.Z. designed the study. B.L. and Y.W. performed the experiments and X.Z. and C.L. assisted with the experiments. B.L., Y.H., X.X. and W.Z. analyzed the data. B.L. and W.Z. wrote the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Liu, B. *et al.* Discovery of a rhamnose utilization pathway and rhamnose-inducible promoters in *Pichia pastoris*. *Sci. Rep.* **6**, 27352; doi: 10.1038/srep27352 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>