Isolation of xylose isomerases by sequence- and function-based screening from a soil metagenomic library

Nádia Skorupa Parachin1,2 and Marie F Gorwa-Grauslund1*

Abstract

Background: Xylose isomerase (XI) catalyses the isomerisation of xylose to xylulose in bacteria and some fungi. Currently, only a limited number of XI genes have been functionally expressed in Saccharomyces cerevisiae, the microorganism of choice for lignocellulosic ethanol production. The objective of the present study was to search for novel XI genes in the vastly diverse microbial habitat present in soil. As the exploitation of microbial diversity is impaired by the ability to cultivate soil microorganisms under standard laboratory conditions, a metagenomic approach, consisting of total DNA extraction from a given environment followed by cloning of DNA into suitable vectors, was undertaken.

Results: A soil metagenomic library was constructed and two screening methods based on protein sequence similarity and enzyme activity were investigated to isolate novel XI encoding genes. These two screening approaches identified the xym1 and xym2 genes, respectively. Sequence and phylogenetic analyses revealed that the genes shared 67% similarity and belonged to different bacterial groups. When xym1 and xym2 were overexpressed in a xylA-deficient Escherichia coli strain, similar growth rates to those in which the Piromyces XI gene was expressed were obtained. However, expression in S. cerevisiae resulted in only one-fourth the growth rate of that obtained for the strain expressing the Piromyces XI gene.

Conclusions: For the first time, the screening of a soil metagenomic library in E. coli resulted in the successful isolation of two active XIs. However, the discrepancy between XI enzyme performance in E. coli and S. cerevisiae suggests that future screening for XI activity from soil should be pursued directly using yeast as a host.

Background

The soil habitat is an immensely diverse environment. One gram of soil may harbour up to ten billion bacteria belonging to more than 4,000 to 7,000 different species [1], although this value varies according to the soil type [2]. However, only 1% of the soil bacterial flora can be cultivated under standard laboratory conditions [3]. Metagenomics, which consists of the extraction, cloning and analysis of the entire genetic material in a given habitat [4,5], has emerged as a tool for assessing the genetic information from uncultivable microorganisms. Metagenomic libraries are a powerful source for discovering new biological activities and have already been successfully used to isolate novel hydrolytic enzymes such as amylases, cellulases and xylanases [6,7].

The screening of metagenomic libraries, which is a critical step for the successful isolation of new and improved biological activities, can be performed by sequence homology or activity-based assays [8]. In sequence-based screening, the isolation of novel proteins is based on homology searches. Although it impedes the discovery of entirely new gene sequences, it is an efficient method of isolating enzymes with different levels of similarity from previously identified genes. For instance, sequence-based screening from the metagenome of a hindgut microbiota of a wood-feeding higher termite identified 700 domains with homology to the glycoside-hydrolase catalytic site [9]. In activity-based assays, metagenomic libraries are constructed in expression vectors to allow for the isolation of biological
activities from entirely new gene sequences that are also successfully expressed in the chosen host.

Activity-based screening methods have been categorized into chromogenic, product detection and growth test assays [10]. Chromogenic assays are commonly used for the detection of hydrolyses from metagenomic libraries [6, 7], whereas a product detection method has been utilized for the identification of enzymes from a soil metagenome producing porphyrin intermediates [11]. Although it is a high-throughput method, growth-based screening has not been frequently described, as it may be difficult to develop and implement a selection protocol that clearly discriminates growth from nongrowth.

Metagenomic libraries have not yet been utilized for the isolation of novel genes involved in xylose catabolism. Xylose is the second most abundant sugar in nature, and the production of cost-efficient lignocellulosic bioethanol requires its complete fermentation [12]. One xylose catabolic pathway is found in naturally xylose-utilizing bacteria [13] and some fungi [14, 15] and includes the isomerisation of xylose to xylulose via xylose isomerase (XI). High ethanol yields from xylose have been reported for Saccharomyces cerevisiae strains harbouring actively expressed XI genes [15-18].

In this study, the possibility of isolating xylose catabolic enzymes from a soil metagenomic library was evaluated. Both sequence-based and growth-based screening were utilized. Each method resulted in the isolation of a novel gene sequence coding for a XI. Restoration of growth on xylose was verified in both Escherichia coli and S. cerevisiae strains by overexpression of the isolated genes in strains lacking the initial xylose catabolic pathway. The advantages and disadvantages of the two screening methods are discussed.

**Results**

**Isolation of XI genes from soil metagenome**

A soil metagenomic library with an estimated size of 1.26 × 10^7 clones was constructed in plasmid pRSETB (Table 1) using DNA extracted from garden compost and screened for XI activity using either sequence analysis or growth assays.

For the sequence-based screening, degenerate primers were designed based on the alignment of amino acid sequences of known XIs from soil microorganisms such as Streptomyces sp. and Rhizobium sp. Amino acid sequences of XI genes whose expression has been attempted in S. cerevisiae, such as the ones from Piyromyces sp. [14] and E. coli [19], were also included (Figure 1). The first polymerase chain reaction (PCR) was performed with primers DF1 and DR2 (Table 2) and with the soil metagenomic library used as a template. A fragment of approximately 750 bp was amplified. The purified fragment was then utilized as the template for a nested PCR with internal primers DF2 and DR1. A single fragment of about 350 bp was obtained, purified from an agarose gel and cloned into the pGEM-T Easy Vector System (Promega, Madison, WI, USA) (Table 1). Sequence analysis revealed that this fragment belonged to a gene coding for a XI.
complete gene sequence was then obtained by PCR utilizing the metagenomic library as the template and primers designed to the sequence of the known 350-bp fragment and in plasmid pRSETB. The gene sequence encoded a 443-amino acid protein with 83% identity with *Sorangium cellulosum* XI and was designated *xym1*.

For the activity screening, the metagenomic library was transferred to the host strain HB101 (Table 1) that has a nonfunctional *xylA*, which allows selection for growth on xylose. After bacterial transformation, cells were plated on defined media supplemented with xylose and incubated at 37°C for four to ten days. Positive (*pRSETB-Xi Piromyces*) and negative (*pRSETB*) controls were also transformed into the screening strain and plated in selective media to validate the screening method. Plasmids were extracted and sequenced from colonies that grew on the selective media. All extracted plasmids contained identical inserted fragments with an insert size of approximately 2.5 kb. Sequencing analysis revealed a complete open reading frame that corresponded to a XI gene named *xym2*. *xym2* encoded a protein of 442 amino acids with 73% similarity to the XI from *Robiginitalea biformata* and 67% identity with *xym1*-encoded XI. A phylogenetic tree constructed using MEGA4 software [20] and 22 different XI amino acid sequences revealed that *xym1*-encoded XI was most similar to XIs from the Proteobacteria, while *xym2*-encoded XI was more similar to the XIs from members of the *Bacteroides* phylum (Figure 2).

**Growth evaluation in E. coli**

The *xym1* and *xym2* genes and the positive control *xylA* gene from *Piromyces* [14] were each cloned into the multicopy vector pRSETB, generating plasmids *pRSETB-Xym1*, *pRSETB-Xym2* and *pRSETB-XiPiromyces* (Table 1). Strain HB101 was transformed with each of the three constructs, as well as the negative control pRSETB, to generate the four strains: TMB2010 (negative control), TMB2011 (*Piromyces xylA*), TMB2012 (*xym1*) and TMB2013 (*xym2*). Growth was evaluated for the four strains in SM3 liquid media with xylose. Growth with xylose as the sole

---

**Table 2 Primers used**

| Primer | Sequence |
|--------|----------|
| DF1    | 5’ AAGCT(T/A/C/G/T)GG(A/C/G/T)(T/A/C/G/T)(T/A/C/G/T)(T/A/C/G/T)AG(A/G)AT 3’ |
| DR2    | 5’ TCTCTC(T/A/C/G/T)TGCTC(T/A/C/G/T)(T/A/C/G/T)GCCGC(T/A/C/G/T)CC(A/C/G/T)AG(A/G)AT 3’ |
| DF2    | 5’ GGT(A/G)(T/A/C/G/T)(T/A/C/G/T)AA(A/C/G/T)(T/A/C/G/T)(T/A/C/G/T)(T/A/C/G/T)AG(A/G)AT 3’ |
| DR1    | 5’ (A/G)(T/A/C/G/T)(T/A/C/G/T)(T/A/C/G/T)(T/A/C/G/T)(T/A/C/G/T)(T/A/C/G/T)(T/A/C/G/T)AG(A/G)AT 3’ |
| pRSETBF| 5’ GGTGAGCACAAAATGGTGCGG 3’ |
| pRSETBR| 5’ GGCGCTTGTAGGACCGCGGATC 3’ |
| XYM1F  | 5’ CCTGACATCTATGAGGCTCTCTTGGCGACAAG 3’ |
| XYM1R  | 5’ TGCTGAGACTTACCGGATCAGCCGTTCAATAG 3’ |
| XYM2F  | 5’ CAGAATTCATGAAATTACCTTACGTTGAGAACAGAAG 3’ |
| XYM2R  | 5’ TGCTGAGACTTAAATACCCGTTACATCAAATTTCTCAATAG 3’ |

*Restriction sites, when present, are underlined.*
carbon source was reestablished for all strains overexpressing any of the XI-encoding genes (Figure 3). No significant differences in growth rates and final Optical density (OD) were observed between the positive control strain, TMB2011, and the strains overexpressing xym1 and xym2 that were isolated from the metagenomic library (Figure 3).

**Growth evaluation in *S. cerevisiae***

XI is a key enzyme for the construction of recombinant *S. cerevisiae* strains for xylose conversion to ethanol. However, despite many efforts [19,21-24], only a few XI genes have been successfully expressed in *S. cerevisiae* at sufficient levels to allow growth on xylose [15,16,25]. Therefore, xym1 and xym2 were overexpressed in *S. cerevisiae* to assess growth in a recombinant *S. cerevisiae* strain lacking the initial xylose catabolic pathway. xym1, xym2 and the xylA gene from *Piromyces sp.* (positive control) were cloned in the multicopy vector p426TEF to generate plasmids p246TEFXiPiromyces, p426TEFXym1 and p426TEFXym2 (Table 1). The *S. cerevisiae* screening strain TMB3044 that has been optimised for efficient pentose utilization, but which lacks the initial conversion step from xylose to xylulose [26], was transformed with the three constructed plasmids in addition to the empty plasmid to generate strains TMB3363 (empty plasmid), TMB3359(XiPiromyces), TMB3364 (xym1) and TMB3365 (xym2). No growth was observed for the negative control. In contrast, overexpression of both xym1 and xym2 enabled growth on xylose in *S. cerevisiae* (Figure 4). However, the strains carrying either xym1 or xym2 grew at a lower growth rate (0.021 hour⁻¹ for both strains) than the positive control harbouring *Piromyces sp* XI (0.069 hour⁻¹ ± 0.006) (Figure 4).

**Evaluation of XI-specific activities in both hosts***

XI-specific activities were compared in cell extracts of *E. coli* and *S. cerevisiae* cells grown in defined medium

---

*Figure 2* Neighbour-joining tree showing the phylogenetic positions of xym1- and xym2-encoded XIs based on conserved amino acids of different XIs. Microorganism sources of XIs were chosen based on sequence similarity and/or active expression in *Saccharomyces cerevisiae*. The phylogenetic tree was constructed with MEGA software using an unweighted group method with arithmetic mean (UFGMA) bootstrapped with 6,000 interactions.
with glucose. For each microorganism, positive (Piro-
myces sp. XI) and negative (empty vector) control strains
were included for comparison with the strains overex-
pressing either xym1 or xym2.

In S. cerevisiae, the background activity measured in
the negative control (0.195 U/mg protein⁻¹) was twice
as high as that measured in E. coli (0.059 U/mg protein⁻¹)
(Figure 5). Strains overexpressing the positive control or
xym1 had about the same specific XI activity in both E.
coli and S. cerevisiae. For the xym2-overexpressing
strain, the specific activity (0.420 U/mg protein⁻¹) was in
the same range as the ranges for the xym1-overexpress-
ing strain and the positive control in E. coli. In con-
trast, the specific activity in the xym2-overexpressing
strain was at the level of the background activity in S.

**Discussion**

In this study, for the first time, two screening methods
were established in parallel for the isolation of novel XI-
encoding genes from a soil metagenomic library. The
sequence-based method using degenerate primers
deduced from the alignment of amino acid sequences of
several XIs from soil microorganisms rapidly resulted in
the isolation of the first XI-encoding gene xym1 from a
metagenomic library. Growth-based screening was also
used for the first time to isolate the XI-encoding gene
xym2 from soil metagenomic library using an E. coli
strain lacking XI activity. The two different methods
resulted in the isolation of two XI-encoding genes that
shared 67% identity at the protein level and could both
complement E. coli and S. cerevisiae strains lacking a
xylene catabolic enzyme.

Some examples where the sequence-based screening
method has already been used include (1) the identifica-
tion of genes responsible for the hydrolytic dehalogene-
tion of 4-chlorobenzoate from a metagenomic library
constructed from a denitrifying, degrading consortium
[27] and (2) the isolation of a polyketide synthase gene
from a metagenome constructed from marine sediment
in the East China Sea [28]. It is a high-throughput
screening method, but one disadvantage is that it selects
for gene sequences with close homology to those
utilized for the design of the degenerate primers. This was notably exemplified by the isolation of closely related related chitinase genes from cultured and uncultured marine bacteria [29].

In contrast, activity-based screening permits the identification of previously unknown gene sequences, as the screening is based on the gene product only. One disadvantage, however, is that activity-based screening that relies on chromogenic or fluorogenic assays often require high-throughput equipment, such as colony pickers, that transfer colonies from plates to the assay system. When possible, it is therefore advisable to develop a growth-based screening method because the method is simple (each library can be plated on a limited number of plates, as very few clones will grow) and enables high-throughput (only the clones displaying the desired activity are able to grow). This method is limited in that a host strain or an experimental setup must be designed that does not enable the host to grow unless the desired activity is present.

Screening by growth has previously been used to isolate genes from a soil metagenomic library conferring Na"(Li")/H" antiporter activity to an antiporter-deficient E. coli host strain [30]. Novel DNA polymerase activities have also been isolated from glacial ice metagenomic libraries using a cold-sensitive E. coli strain that harboured a mutation in the DNA polymerase domain generating lethality at temperatures below 20°C [31]. Finally, a novel D-3-hydroxybutyrate short-chain dehydrogenase with less than 35% similarity to known proteins with similar activity has been isolated from soil metagenomic libraries using a mutant of Sinorhizobium meliloti that was unable to synthesize D-3-hydroxybutyrate [32]. In the present study, xym1- and xym2-encoded XIs displayed strong similarity in amino acid sequence, although they were isolated by using different methodologies. This may be explained by the fact that most of the XI sequences that were used to identify homologous regions for the sequence-based screening originated from organisms that were known to be present in soil.

Screening for XI-encoding genes by growth versus nongrowth was successfully implemented in E. coli, since the expression of xym1 and xym2 genes could restore growth in xylose minimal media with the same growth rate observed for the positive control overexpressing the Piromyces XI gene. However, growth was restored when the same genes were transferred to the heterologous host S. cerevisiae, where the activity is required for xylose fermentation to ethanol, the growth rate was four times lower than that observed for Piromyces XI. Expression of XI-encoding genes in S. cerevisiae is a complex issue [33]. The first one to be expressed at high levels in S. cerevisiae was Piromyces XI [18], and it occurred after many unsuccessful attempts and limited success with several other bacterial and plant XIs [19,21-24]. Successful expression of XI in S. cerevisiae and its growth rate on xylose are not always correlated to sequence similarity or enzymatic activity (Table 3). In our case, xym1-encoded XI had the same XI activity as Piromyces sp XI, but the growth rate was fourfold lower. In contrast, the specific activity of the xym2-encoded XI was in the range of the background activity, but the growth rate was identical to the strain expressing xym1. The growth rate was also in the same range when the Xanthomonas campestris XI gene was expressed, although the reported activity range was about 20 times lower than that for the xym1-overexpressing strain (Table 3). Finally, a high activity level and growth rate were achieved when overexpressing a codon-optimised XI gene from Clostridium phytofermentans that shared only 54% similarity with Piromyces sp. at the protein level (Table 3) [16]. Differences in kinetic properties may also explain the discrepancies among all these results. For instance, the affinity for xylose may be much higher in xym2 than in xym1-encoded XI, so that the lower rate of conversion would be compensated for by a higher affinity for the substrate. Alternatively, the novel XI may be more inhibited by xyitol than the one from Piromyces sp. or C. phytofermentans. We were not able to measure K_m or K_i for the new XIs, because the activities were below the detection limit. Transport may also explain differences in behaviour between S. cerevisiae and E. coli, since active xylose transporters are present in E. coli [34,35] but missing in S. cerevisiae.

E. coli has been the preferred choice as the host for screening metagenomic libraries [6] because numerous molecular tools are available and E. coli genetics are well elucidated. In our case, E. coli had the additional advantage that it is naturally able to utilize xylose, allowing the isolation of the desired activity without any
other limitation up- or downstream of XI. The use of only one host, however, may limit the exploitation of microbial diversity for a given habitat due to, for instance, differences in codon utilization and/or mRNA and protein stability among different hosts. *E. coli* was, for instance, estimated to express only 40% of genes from diverse microbial origin [36]. In a previous study where growth-based screening was utilized for the isolation of novel D-3-hydroxybutyrate short-chain dehydrogenases from the soil metagenomic library, 25 positive clones were initially isolated when the library was screened in *Sinorhizobium meliloti*; however, only one clone complemented the same phenotype in *E. coli* [32].

Screening of a soil metagenomic library performed in six different bacterial hosts identified only one case where the molecule-producing clone could be isolated using two different hosts [37]. Further screening of the soil metagenomic library described in this study should be performed in *S. cerevisiae* to investigate more suitable XIs for ethanol production from xylose. This is complicated by the fact that, whereas in bacterial metagenomic libraries large fragments are cloned and theoretically all genes within the fragment are transcribed and translated, only the closest gene to the methylated cap on the 5’ terminus of the mRNA is translated in yeast [38]. Construction of future libraries for use in yeast will have to combine the cloning of small fragments corresponding to single gene size with an efficient transformation protocol to accommodate the larger colony numbers required to represent the same soil population. Alternatively, construction and screening of metagenomic cDNA libraries may be attempted; however, neither of these strategies has yet been reported.

**Conclusions**

In the present study, a soil metagenomic library was successfully constructed and screened for the first time by using sequence- and growth-based methods to isolate genes conferring xylose isomerase activity. It enabled the isolation of *xym1* and *xym2* genes, respectively, that shared 67% identity at the protein level. However, the limited growth rates in *S. cerevisiae* highlight the importance of screening for XI activity in this host instead.

**Methods**

**Plasmids and strains**

Plasmids and strains used in this study are listed in Table 1. Strains were stored as recommended by the manufacturer or as 20% glycerol stocks in liquid media at -80°C.
Cultivation conditions

*E. coli* strains were plated in Luria broth (LB) media plates (10 g/L Tryptone, 5 g/L yeast extract and 10 g/L NaCl, agar 15 g/L) and grown at 37°C overnight. Antibiotics were added when necessary to a final concentration of 30 μg/mL kanamycin and 50 μg/mL ampicillin and streptomycin. Single *E. coli* colonies were inoculated in LB media (10 g/L Bacto Tryptone, 5 g/L yeast extract and 10 g/L NaCl) for approximately 12 hours at 37°C in an incubator with constant agitation at 180 rpm (Boyle Medical AB, Stockholm, Sweden).

For the metagenomic library screening, cells were plated in defined SM3 media [39] supplemented with 11 g/L xylose, 15 g/L agar, 1 mM isopropyl-β-D-thiogalactoside (IPTG) and 50 μg/ml ampicillin.

*S. cerevisiae* strains were grown on yeast nitrogen base (YNB) media plates (6.7 g/L Difco YNB without amino acids; Becton Dickinson, Sparks, MD, USA) supplemented with 20 g/L glucose and 20 g/L agar.

For growth kinetics on xylose, *E. coli* recombinant strains were pregrown in 20 mL of LB media at 180 rpm and 30°C until the end of the exponential phase. Cells were washed in distilled water and transferred to 500-mL Erlenmeyer flasks containing 50 mL of SM3 media supplemented with 30 g/L xylose and 1 mM IPTG at a starting OD₆₂₀nm = 0.1.

*S. cerevisiae* strains were pregrown in 500-mL Erlenmeyer flasks containing 50 mL of YNB medium supplemented with 20 g/L glucose at 180 rpm and 30°C for approximately 12 hours. Cells were washed in distilled water and inoculated at OD₆₂₀nm = 0.2 in 100 mL of YNB medium supplemented with 50 g/L xylose in 1-L Erlenmeyer flasks and grown at 180 rpm and 30°C.

Standard molecular procedures

Standard DNA manipulation techniques were used [40]. Restriction and ligation enzymes were purchased from Fermentas (St. Leon-Rot, Germany). Primers were purchased from Eurofins MWG Operon (Ebersberg, Germany). DNA extraction from agarose gels and purification of PCR products were performed using the QIAquick extraction kit (Qiagen, Hilder, Germany). DNA extraction was performed using the PowerSoil DNA Isolation Kit (MoBio, Carlsbad, CA, USA). Isolated DNA was digested with *BamHI* and *MboI*. Agarose gel electrophoresis was performed for purification of DNA fragments ranging between 2.0 and 6.0 kb to generate a small insert library with an average size of 4.0 kb. Vector pRSETB was digested with *BamHI* and treated with alkaline phosphatase to prevent plasmid self-ligation. DNA fragments were ligated into pRSETB. Bacterial transformation was performed with ElectroTen-Blue competent cells (Stratagene, La Jolla, CA, USA) according to the manufacturer’s instructions. The library size was estimated on the basis of the total number of *E. coli* clones minus the ones with empty plasmid.

Cloning and expression of *xym1*, *xym2* and XI from *Piromyces sp*

The isolated isomerase genes were cloned both into *E. coli* and *S. cerevisiae* expression vectors. The *xym1* gene was amplified with primers XYM1F and XYM1R, which contain restriction sites for *BamHI* and *SalI*, respectively (Table 2). After PCR, both the amplified gene and plasmids pRSETB and p426TEF were digested. Plasmids with inserts obtained after ligation and bacterial transformation were sequenced and named pRSETB-Xym1 and p426TEF-Xym1, respectively (Table 1). Cloning of the *xym2* gene followed the same procedure as for *xym1*, with the exception that *EcoRI* and *SalI* restriction sites were used. *xym1* and *xym2* sequences will be publicly available when the corresponding patent application is published.

A codon-optimised version of the XI-encoding gene from *Piromyces sp*. was amplified by PCR utilizing plasmid YeplacHXT-XI as template [26]: forward primer 5′ GGAGGATCCATGGAATAATTTTCCACAATTCTGCTCCTTGCAAA 3′ and reverse primer 5′ TTGAAGATTCTTACTG ATACATTTGCAAACATAGCTTTGGAT3′. Restriction sites for *BamHI* and *EcoRI* are underlined in the forward and reverse primers, respectively. Plasmid p426TEF [43], pRSETB (Invitrogen, Carlsbad, CA, USA) and the PCR product were digested with *BamHI* and *EcoRI* and then ligated. Clones with inserts were sequenced prior to yeast transformation.

Enzyme activities

*S. cerevisiae* strains were grown in YNB medium supplemented with 20 g/L glucose at 30°C to the early stationary phase. *E. coli* strains were pregrown in LB media supplemented with 100 μg/ml ampicillin for about 5 hours. After that, cells were washed with water and inoculated for an initial OD₆₂₀nm = 0.2 in SM3 medium supplemented with 10 g/L glucose and 1 mM IPTG at 37°C overnight.

Yeast and bacteria strains were harvested by centrifugation at 5,000 x g for 5 minutes and washed once with DNA extraction was performed using the PowerSoil DNA Isolation Kit (MoBio, Carlsbad, CA, USA). Isolated DNA was digested with *BamHI* and *MboI*. Agarose gel electrophoresis was performed for purification of DNA fragments ranging between 2.0 and 6.0 kb to generate a small insert library with an average size of 4.0 kb. Vector pRSETB was digested with *BamHI* and treated with alkaline phosphatase to prevent plasmid self-ligation. DNA fragments were ligated into pRSETB. Bacterial transformation was performed with ElectroTen-Blue competent cells (Stratagene, La Jolla, CA, USA) according to the manufacturer’s instructions. The library size was estimated on the basis of the total number of *E. coli* clones minus the ones with empty plasmid.

Cloning and expression of *xym1*, *xym2* and XI from *Piromyces sp*

The isolated isomerase genes were cloned both into *E. coli* and *S. cerevisiae* expression vectors. The *xym1* gene was amplified with primers XYM1F and XYM1R, which contain restriction sites for *BamHI* and *SalI*, respectively (Table 2). After PCR, both the amplified gene and plasmids pRSETB and p426TEF were digested. Plasmids with inserts obtained after ligation and bacterial transformation were sequenced and named pRSETB-Xym1 and p426TEF-Xym1, respectively (Table 1). Cloning of the *xym2* gene followed the same procedure as for *xym1*, with the exception that *EcoRI* and *SalI* restriction sites were used. *xym1* and *xym2* sequences will be publicly available when the corresponding patent application is published.

A codon-optimised version of the XI-encoding gene from *Piromyces sp*. was amplified by PCR utilizing plasmid YeplacHXT-XI as template [26]: forward primer 5′ GGAGGATCCATGGAATAATTTTCCACAATTCTGCTCCTTGCAAA 3′ and reverse primer 5′ TTGAAGATTCTTACTG ATACATTTGCAAACATAGCTTTGGAT3′. Restriction sites for *BamHI* and *EcoRI* are underlined in the forward and reverse primers, respectively. Plasmid p426TEF [43], pRSETB (Invitrogen, Carlsbad, CA, USA) and the PCR product were digested with *BamHI* and *EcoRI* and then ligated. Clones with inserts were sequenced prior to yeast transformation.

Enzyme activities

*S. cerevisiae* strains were grown in YNB medium supplemented with 20 g/L glucose at 30°C to the early stationary phase. *E. coli* strains were pregrown in LB media supplemented with 100 μg/ml ampicillin for about 5 hours. After that, cells were washed with water and inoculated for an initial OD₆₂₀nm = 0.2 in SM3 medium supplemented with 10 g/L glucose and 1 mM IPTG at 37°C overnight.

Yeast and bacteria strains were harvested by centrifugation at 5,000 x g for 5 minutes and washed once with
distilled water. Wet cells were suspended (0.5 mg/mL) in Y-per yeast or B-per bacteria extraction protein reagent (Pierce Biotechnology, Rockford, IL, USA) in a 2-mL Eppendorf tube and incubated on a turning table at room temperature for 50 minutes. Cell debris was spun down for 5 minutes at 16,100 × g (Z160M table centrifuge; HERMLE Labortechnik, Wehingen, Germany). Protein concentration was determined using Coomassie Protein Assay Reagent (Pierce Biotechnology). Bovine serum albumin was used to determine the standard curve.

XI activity was determined at 30°C using sorbitol dehydrogenase as previously described [18]. Assays were performed in a U-2000 spectrophotometer (Hitachi, Tokyo, Japan). Enzyme assays were performed in three biological replicates.

Acknowledgements
We thank Prof. Barbel Hahn-Hägerdal for her valuable comments, Catherine Paul for critically reading the manuscript and Birgit Johanson for providing fresh soil for DNA extraction. This work was financed by the Swedish Research Council (VR).

Author details
1Department of Applied Microbiology, Center for Chemistry and Chemical Engineering, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden.
2Present affiliation - Laboratorio de Biologia Molecular, Universidade de Brasilia, 70910-900 Brasilia (DF), Brazil.

Authors’ contributions
NSP planned and carried out the experimental work and wrote the manuscript. MFGG participated in the study design and its coordination and critically read and commented on the manuscript. Both authors reviewed and approved the final version of the manuscript.

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

Received: 11 January 2011 Accepted: 5 May 2011 Published: 5 May 2011

References
1. Rosselló-Mora R, Amann R. The species concept for prokaryotes. FEMS Microbiol Rev 2001, 25:39-67.
2. Torsvik V, Øvreås L. Prokaryotic diversity: magnitude, dynamics, and controlling factors. Science 2002, 296:1646-1648.
3. Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM. Microbial diversity and function in soil: from genes to biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Biochemistry 1998, 37:7515-7518.
4. Kuyper M, Harhangi HR, Stave AK, Winkler AA, Jetten MS, de Laat WT, den Ridder JJ, Op den Camp HJ, van Dijken JP, Pronk JT. High-level functional expression of a fungal xylose isomerase: the key to efficient ethanol fermentation of xylose by Saccharomyces cerevisiae? FEMS Yeast Res 2003, 3:69-78.
5. Sarthy AV, McConaughy BL, Lobo Z, Sundstrom JA, Furlong CE, Hall BD. Expression of the Escherichia coli xylose isomerase gene in Saccharomyces cerevisiae. Appl Environ Microbiol 2009, 75:2304-2311.
6. Karhumaa K, Garcia-Sanchez R, Hahn-Hägerdal B, Gorwa-Grauslund MF. Comparison of the xylose reductase-xyitol dehydrogenase and the xylose isomerase pathways for xylose fermentation by recombinant Saccharomyces cerevisiae. Microb Cell Fact 2007, 65.
7. Jiao YL, Wang LH, Dong XY, Chen YF, Zong Y, Gao Y, Ren N, Guo AY. Identification of genes coding for hydrolytic dehalogenation in the metagenome derived from a denitrifying 4-chlorobenzoate degrading consortium. FEMS Microbiol Lett 2008, 281:203-209.
8. Jiao YL, Wang LH, Dong XY, Chen YF, Zong Y, Gao Y, Ren N, Guo AY, Zhang XQ, Jiao BH. Isolation of new polyketide synthase gene fragments and a partial gene cluster from East China Sea and function analysis of a new acyltransferase. Appl Biochem Biotechnol 2008, 149:655-664.
9. Cottrell MT, Wood DN, Yu L, Kirchman DL. Selected chitinase genes in cultured and uncultured marine bacteria in the α- and γ-subclasses of the proteobacteria. Appl Environ Microbiol 2000, 66:1195-1201.

Metagenomic and functional analysis of hindgut microbtiota of a wood-feeding higher termite. Nature 2007, 450:560-565.
30. Majerník A, Gottschalk G, Daniel R. Screening of environmental DNA libraries for the presence of genes conferring Na\(^+\)(Li\(^+\))/H\(^+\) antiporter activity on *Escherichia coli*: characterization of the recovered genes and the corresponding gene products. *J Bacteriol* 2001, 183:6645-6653.

31. Simon C, Hesath J, Rockstroh S, Daniel R. Rapid identification of genes encoding DNA polymerases by function-based screening of metagenomic libraries derived from glacial ice. *Appl Environ Microbiol* 2009, 75:2964-2968.

32. Wang C, Meek DJ, Panchal P, Boruva N, Archibald FS, Driscoll BT, Charles TC. Isolation of poly-3-hydroxybutyrate metabolism genes from complex microbial communities by phenotypic complementation of bacterial mutants. *Appl Environ Microbiol* 2006, 72:384-391.

33. Hahn-Hägerdal B, Karhumaa K, Jeppsson M, Gorwa-Grauslund MF. Metabolic engineering for pentose utilization in *Saccharomyces cerevisiae*. Adv Biochem Eng Biotechnol 2007, 108:147-177.

34. Davis EC, Henderson PJ. The cloning and DNA sequence of the gene xylE for xylose-proton symport in *Escherichia coli* K12. *J Biol Chem* 1987, 262:13928-13932.

35. Kurose N, Watanabe K, Kimura A. Nucleotide sequence of the gene responsible for D-xylose uptake in *Escherichia coli*. Nucleic Acids Res 1986, 14:7115-7123.

36. Gabor EM, Alkema WB, Janssen DB. Quantifying the accessibility of the metagenome by random expression cloning techniques. *Environ Microbiol* 2004, 6:879-886.

37. Craig JW, Chang FY, Kim JH, Obiajulu SC, Brady SF. Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental DNA libraries in diverse proteobacteria. *Appl Environ Microbiol* 2010, 76:1633-1641.

38. Lewin B. *Genes II* New York: Oxford University Press; 2000.

39. Rothen SA, Sauer M, Sonnleitner B, Wirth B. Growth Characteristics of *Escherichia coli* HB101[pGEc47] on defined medium. *Biotechnol Boeng* 1998, 58:92-100.

40. Sambrook J, Fritch E, Maniatis T. Molecular Cloning: A Laboratory Manual. 2 edition. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 1989.

41. Sambrook J, Russell DW. Molecular Cloning: A Laboratory Manual. 3 edition. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 2001.

42. Gietz RD, Schiestl RH, Willems AR, Woods RA. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* 1995, 11:355-360.

43. Mumberg D, Muller R, Funk M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 1995, 156:119-122.

44. Ejiofor CG. Molecular tools for improving xylose fermentation in xylose isomerase expressing yeasts. *MSc thesis* Lund University, Department of Applied Microbiology; 2004.

45. Winkler AA, Kuyper SM, De Laat WT, Van Dijken JP, Pronk JT. Metabolic engineering for pentose utilization in *Saccharomyces cerevisiae*: expression of genes involved in pentose metabolism. PhD thesis Lund University, Department of Applied Microbiology; 1995.

46. Hallborn J. *Biotechnol Bioeng* 1987, 26:147-177.

47. Amore R, Wilhelm M, Hollinberg CP. The fermentation of xylose: an analysis of the expression of *Bacillus* and *Actinoplanes* xylose isomerase genes in yeast. *Appl Microbiol Biotechnol* 1989, 30:351-357.

48. Lönn A, Tarff-Bjerre KL, Cordero Otero RR, van Zyl WH, Hahn-Hägerdal B. Xylose isomerase activity influences xylose fermentation with recombinant *Saccharomyces cerevisiae* strains expressing mutated xylA from *Thermus thermophilus*. *Enzyme Microb Technol* 2003, 32:567-573.

Cite this article as: Parachin and Gorwa-Grauslund: Isolation of xylose isomerases by sequence- and function-based screening from a soil metagenomic library. *Biotechnology for Biofuels* 2011 4:9.

doi:10.1186/1754-6834-4-9

Submit your next manuscript to *BioMed Central* and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit