VERSITA

CELLULAR & MOLECULAR BIOLOGY LETTERS
Volume 11 (2006) pp 348 - 359
http://www.cmbl.org.pl

DOI: 10.2478/s11658-006-0029-z

Received: 23 November 2005
Revised form accepted: 15 May 2006

GENOTYPING OF RECOMBINANT *Pichia pastoris* STRAINS

JOSÉ M. VIADER-SALVADÓ*, EDDY L. CAB-BARRERA, LUIS J. GALÁN-WONG and MARTHA GUERRERO-OLAZARÁN
Instituto de Biotecnología, Facultad de Ciencias Biológicas, Universidad Autónoma de Nuevo León. Ave. Pedro de Alba s/n, Col. Ciudad Universitaria, 66450 San Nicolás de los Garza, N.L., Mexico

Abstract: A simplified amplified-fragment length polymorphism (AFLP) method was used to genotype *Pichia pastoris* strains obtained by transformation of *P. pastoris* strain GS115 with a single integration vector. A total of 14 transformants and 3 control strains were analyzed, which generated 16 different band patterns. A clonal variation was obtained after the transformation process due to genetic differences generated during the transformation event of the host strain. Furthermore, the cluster analysis showed that the transformants with lesser genetic differences with respect to the *P. pastoris* host strain are the recombinant strains with the highest level of recombinant protein production.

Key words: *Pichia pastoris*, Amplified-fragment length polymorphism (AFLP), Clonal variation

* Corresponding author: e-mail: jviader@fcb.uanl.mx

Abbreviations used: AFLP – amplified-fragment length polymorphism; *AOX1* – alcohol oxidase 1 gene; BMG – growth medium: 100 mM potassium phosphate, pH 6.0, 0.4 mg/l biotin, 1.34% YNB, 1% glycerol; BMM – expression medium: 100 mM potassium phosphate, pH 6.0, 0.4 mg/l biotin, 1.34% YNB, 0.75% (v/v) methanol; DMT-off – without dimethoxytrityl; n.d. – not determined; PCR – polymerase chain reaction; RDB – 1 M sorbitol, 1.34% YNB, 2% dextrose, 0.4 mg/l biotin, 0.005% amino acids without histidine; r^2 – coefficient of determination; RP – recombinant protein; SDS – sodium dodecyl sulphate; TE buffer – 10 mM Tris-Cl, 1 mM EDTA, pH 8.0; TP – total protein; XA-1 and XA-2 – *Xho*I site cohesive-ended complementary oligonucleotides to generate the *Xho*I restriction fragment adapter; XP-G – PCR primer used in the AFLP reactions directed at the adapter tail and with a G as selective nucleotide at the 3'-end; YNB – yeast nitrogen base; YPD – 1% yeast extract, 2% peptone, 2% dextrose.

INTRODUCTION

The expression system of *Pichia pastoris* offers many advantages that are comparable to those of higher organisms, two of the most important being the ability to perform post-translational modifications and the capacity to yield the correct protein folding. In addition, its manipulation for biotechnological purposes is simpler and less expensive compared to other eukaryotic hosts used for this purpose, and higher expression levels are generally obtained [1, 2].

The introduction of DNA into the *P. pastoris* genome occurs by homologous recombination between the transforming vector and the *P. pastoris* genome. Thus, the integration of vector sequences is carried out at predetermined positions in the yeast genome [3]. In theory, these integration events are controlled, though after *P. pastoris* transformation with an expression vector, a clonal variation in the heterologous gene expression has been observed [4-6]. Cultures of many transformant strains need to be carried out to select the strain that produces the highest amount of recombinant protein. Although the presence of this clonal variation has been suggested to be due to differences in the number of integrated copies of the expression cassette in the yeast genome [7, 8], for some proteins [9, 10], a single copy of the expression cassette was sufficient for the optimal production of recombinant protein, indicating that the clonal variation is not only due to the copy number of the expression cassette.

If the clonal variation obtained after the transformation of *P. pastoris* strains could be evaluated through genetic characterization of the resulting colonies, similar colonies could be classified to facilitate the strain selection. Thus, the evaluation of the gene expression levels could be done on a smaller number of transformant strains. In addition, with the genotyping analysis, each strain could be characterized with a high degree of confidence for subsequent applications. This genotypic characterization could also follow the genetic stability of selected strains along multiple subcultures.

Among the diverse possibilities of genotyping microorganisms, techniques based on the polymerase chain reaction (PCR) are especially useful since they use a very small amount of genomic DNA and yield a large number of highly reproducible polymorphic markers. One of the most promising techniques is amplified-fragment length polymorphism (AFLP) [11], originally developed to identify plant varieties and improve cultivars [11, 12]. In microbiology, AFLP has been employed for typing both fungi and bacteria [13, 14]. Recently, it was applied to study the genetic relationships between varieties of *Pichia kluyveri* collected from different locations around the world [15].

In this study, a simplified AFLP method was used as a genotyping method for several strains of *Pichia pastoris* obtained by the transformation of the *P. pastoris* host strain GS115 (*his4*) with an integration vector. This made it possible to study the genetic differences between closely related transformant strains.

MATERIALS AND METHODS

Transformation into the *Pichia pastoris* strain GS115

The *P. pastoris* host strain GS115 (*his4*), from Invitrogen (San Diego, CA), was transformed by the spheroplasting method as described elsewhere [16, 17] with a *Sal*I linearized plasmid. This plasmid was constructed previously in our laboratory (unpublished results) from the expression vector pPIC9K (Invitrogen, San Diego, CA). It carries the kanamycin-resistance gene, which confers dose-dependent resistance to the antibiotic G418 in *P. pastoris* and is used for selecting multicopy transformants. Spontaneous generation of multiple insertion events can be identified by the level of resistance to G418. Additionally, the plasmid harbors the *HIS4* gene, which complements *his4* in the host, and a 528-bp cDNA that codes for a 20-kDa protein under the control of the alcohol oxidase 1 (*AOX1*) promoter and that is fused to the α-factor signal sequence of *Saccharomyces cerevisiae* to direct the secretion of the recombinant protein into the growth medium. Therefore, the recombinant protein was expected to be secreted into the culture medium. The recombinant strains were selected for their ability to grow on histidine-deficient medium (RDB-agar plates) at 30ºC until colonies appeared (His^+ selection). A total of 200 colonies were obtained, of which 14 colonies were randomly selected and grown a second time in RDB-agar.

Characterization of transformants

The transformants were grown in 10 ml of YPD broth for 24 h at 30ºC, and cells were harvested by centrifugation at 16,000 g for 5 min. For each of the 14 selected recombinant strains, the level of resistance to G418 was evaluated by growing each on YPD-agar plates with increasing doses of the antibiotic [7]. The integration of the expression cassette into the *P. pastoris* genome was evaluated by PCR using DNA isolated from transformants and the 5' and 3' *AOX1* primers (Invitrogen) directed to the *AOX1* promoter and terminator, as previously described [18]. Amplified products were resolved by electrophoresis in a 0.6% (w/v) agarose gel and stained with ethidium bromide (2 μg/ml).

The transformed strains were also grown in 25 ml of BMG medium in 250-ml shake flasks, and subsequently grown in BMM medium for 140 h following the general recommendations from Invitrogen [16]. The protein concentrations in the cell-free culture medium were determined by the Bradford protein assay. In addition, 1 ml cell-free culture medium from each grown strain was dialyzed with a 10-kDa cut-off membrane and dried in a vacuum concentrator. The relative abundances of the 20-kDa recombinant protein, with respect to total protein content, were determined by densitometric analysis of the Coomassie blue-stained protein bands, previously separated on discontinuous 6-12% (w/v) SDS-polyacrylamide gels, using an EDAS 290 system and the Kodak Digital Science 1D program (Eastman Kodak Company, Rochester, New York). Recombinant protein levels in the cell-free culture medium from each grown

strain were estimated with both data sets (the total protein concentration and the abundances of the 20-kDa recombinant protein).

AFLP reactions

The XA-1 (GTAGACTGCGTACATGCA) and XA-2 (TCGATGCATGT ACGCAGT) cohesive-ended complementary oligonucleotides and the XP-G (TGCGTACATGCATCGAGG) PCR primer used in the AFLP reactions were designed specifically for this study, so that the procedure could be carried out using only one restriction enzyme, one double strand adapter, and one PCR primer with one additional G at the 3'-end. These oligonucleotides were DMT-off synthesized (Operon, Alameda, CA), dissolved in water, and stored at -20ºC until used, without prior purification. *Xho*I was selected as the restriction enzyme in the first step of the AFLP reactions, since the cloning vector used to transform the *P. pastoris* host strain GS115 contains two restriction sites for this enzyme.

For the isolation of genomic DNA from yeast, the 14 randomly selected *P. pastoris* transformants and three control strains (GS115 and two *P. pastoris* transformed with *Sal*I-linearized pPIC9K vector without any expression cassette) were grown into 4 ml YPD medium for 24 h at 30ºC with shaking. Cells were harvested by centrifugation at 16,000 g for 5 min, and the cell pellets were resuspended in 200 µl of lysis buffer (2% Triton X-100, 1% SDS, 100 mM NaCl, 1 mM EDTA, 10 mM Tris-HCl, pH 8.0), 500 µl of Tris-HCl saturated phenol, and 100 µl of chloroform/isoamyl alcohol (24:1). The suspensions were vortexed for 5 min, and 200 µl TE buffer (10 mM Tris-HCl and 1 mM EDTA, pH 8.0) was added. The phases were separated by centrifugation at 16,000 g for 10 min, and the aqueous phases were ethanol-precipitated. The DNA pellets were dried, resuspended in 48 µl of TE buffer and 2 µl of 2 µg/µl RNase, and quantified by densitometry of the agarose gels stained with 2 µg/ml ethidium bromide using the EDAS 290 system and the Kodak Digital Science 1D program (Eastman Kodak Company, Rochester, New York).

The first step of the AFLP assay was carried out by digesting 200 ng of genomic DNA with 10 U of *Xho*I (Roche Applied Science, Indianapolis, IN), 2.5 µl of 10X buffer (buffer H, Roche Applied Science), and sterile water in a 25-µl reaction volume for 2 h at 37ºC. The double strand adapter was prepared by mixing XA-1 and XA-2 oligonucleotides at 25 µM each, containing 90 mM NaCl, 10 mM Tris-HCl, and 0.1 mM EDTA (pH 8.0). This mixture was heated at 90ºC for 5 min, cooled slowly to room temperature, and then stored at -20ºC until used in the ligation step. Fifteen µl of *Xho*I-digested DNA were ligated with the adapter at a final concentration of 2.5 µM and 1 U of T4 DNA ligase (Promega, Madison, WI), 2.5 µl of 10X reaction buffer (Promega, Madison, WI) in a 25-µl reaction volume for 17 h at 12ºC. Then, the T4 DNA ligase was thermally inactivated at 65ºC for 10 min. Forty µl of sterile water, 50 µl of 7.5 M NH_4AcO, and 300 µl of 95% (v/v) ethanol were added to each reaction tube. This mixture was centrifuged at 10,000 rpm for 8 min, and the supernatant

removed. The pellet was washed with 250 µl of 75% (v/v) ethanol, air dried for 10 min, dissolved in 25 µl of 10 mM Tris-HCl, 0.1 mM EDTA (pH 8), and stored at -20ºC until use. The PCR step was performed in a PCR Sprint thermal cycler (Hybaid, Middlesex, UK) in a 25-µl reaction volume containing 4.2 µl of dissolved ligation mixture, 10X PCR buffer (Promega, Madison, WI), 1.5 mM $MgCl_2$, 200 µM dNTPs, 6.0 µM of XP-G primer, sterile water, and 2 U of *Taq* DNA polymerase (Promega, Madison, WI). The touchdown PCR technique [19], in a three-step thermal program, was used as follows: 72ºC for 2 min; a 16-cycle amplification of 94ºC for 20 s, 65ºC for 30 s, decreasing by 1ºC each cycle, and 72ºC for 2 min, with an initial denaturation step of 94ºC for 2 min; and a 19-cycle amplification of 94ºC for 30 s, 50ºC for 30 s, and 72ºC for 2 min, with a final elongation step of 60ºC for 30 min. After the PCR assays, 4 µl of the reaction mixture was analyzed in a 2.5% (w/v) agarose electrophoresis gel stained with 2 µg/ml ethidium bromide using a mini-single cell (Fotodyne, Hartland, WI) at 100 volts for 60 to 70 min. The molecular sizes and intensities of the visualized bands were determined using the EDAS 290 system and the Kodak Digital Science 1D program (Eastman Kodak Company, Rochester, New York).

Numerical analysis of the banding patterns

The banding patterns obtained from each analyzed sample were manually scored as binary data with band presence as "1" and band absence as "0", and then expressed as a vector of a 24-dimensional space (the total number of different bands obtained). The similarity indexes of all the possible pairwise comparisons between pairs of the banding patterns of the transformed *P. pastoris* strains (91 pairs) were calculated through the cosine distances between the feature vectors, defined by the ratio of the scalar vector product of each zero-one vector pair and the product of the length of the same vector pair. A similarity matrix containing the similarity indexes, and the Neighbor and Drawgram programs of the Phylip package of programs [20] were used to construct a distance tree dendrogram-type plot by the unweighted pair group method, with arithmetic mean (UPGMA) of clustering [21] calculated according to the banding pattern similarity.

RESULTS

PCR integration analysis of the DNA isolated from all the randomly selected transformants showed two bands: 2107 and 1016 bp (Fig. 1, lanes 3 to 9), which were derived from the *AOX1 locus* (Fig. 1, lane 10) and from the integrated expression cassette (Fig. 1, lane 11) indicating the correct integration of the expression cassette into the *P. pastoris* genome, as well as the integrity of the *AOX1* structural gene. These results confirm the Mut^+ genotype of these recombinant strains (methanol utilization-positive, because the endogenous *AOX1* gene is intact). All the strains were classified according to their resistance to G418 (i.e. resistant to 0.25, 0.75, 1.00, 1.50 or 3.00 mg/ml G418). SDS-polyacrylamide gels of dialyzed cell-free culture medium from recombinant

P. pastoris strains grown in methanol clearly showed a predominant protein band with the expected molecular weight of 20 kDa (Fig. 2, right). The total protein (TP) concentrations of transformants grown in BMM shake-flask cultures ranged from 17.6 to 34.2 mg/l of cell-free culture medium, with the recombinant protein comprising 21 to 52% of the total protein, corresponding to 5.1 to 17.6 mg/l (Fig. 4).

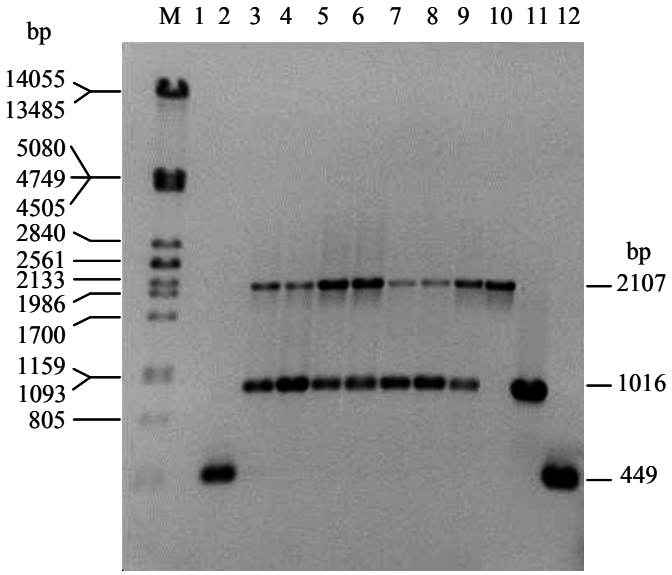


Fig. 1. Results of the PCR integration analysis of the expression cassette in recombinant *P. pastoris* strains, on a 0.6% (w/v) agarose electrophoresis gel. The numbers on the sides are in base pairs. Lane M, molecular weight standards lambda-*Pst*I; lane 1, negative control of PCR (amplification without target DNA); lanes 2 and 12, PCR product of pPIC9K (449 bp) as a control; lanes 3 to 9, PCR product of genomic DNA from seven recombinant *P. pastoris* strains; lane 10, PCR product of GS115 genomic DNA (2107 bp); lane 11, PCR product of the pPIC9K plasmid harboring the expression cassette (1016 bp) as a control.

The reproducibility of the AFLP genotyping method was evaluated by carrying out the complete assay three times with three DNA samples extracted from the same strain (*P. pastoris* GS115). In each of the three assays to test reproducibility, 8 bands, of 1470, 869, 830, 595, 467, 445, 393 and 260 bp, were obtained with a coefficient of variation (ratio of the standard deviation to the absolute value of the mean times 100) of relative intensities of less than 10%.

Fig. 2 (left) shows an agarose gel with the AFLP-band patterns of five of the analyzed strains. The designed oligonucleotides and PCR primer rendered with the strains analyzed a total of 24 different bands from 1470 to 61 bp, in combination, generating 16 different band patterns. Only seven (29%) of the bands were present in more than 50% of the analyzed strains. A 260-bp band

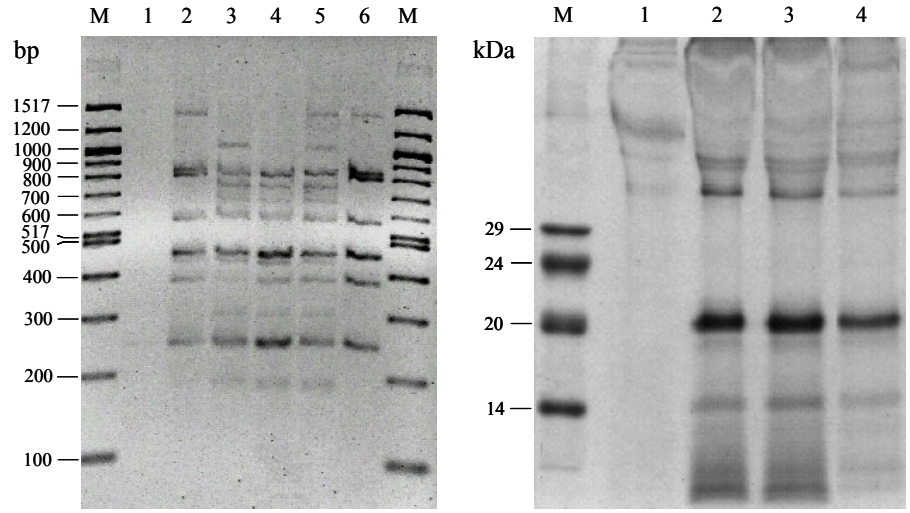


Fig. 2. Agarose gel of the AFLP assays (left) and SDS-polyacrylamide gel of the dialyzed cell-free culture medium (right) from several recombinant *P. pastoris* strains. Left: lane M, molecular weight marker; lane 1, negative AFLP control (amplification without target DNA); lane 2, AFLP products of *P. pastoris* host strain GS115; lanes 3 to 6, several AFLP products of genomic DNA from recombinant *P. pastoris* strains. Right: lane M, molecular weight marker; lane 1, cell-free culture medium from pPIC9K-transformant (without any expression cassette) grown in methanol (negative control); lanes 2 to 4, cell-free culture medium from three different recombinant *P. pastoris* strains grown in methanol.

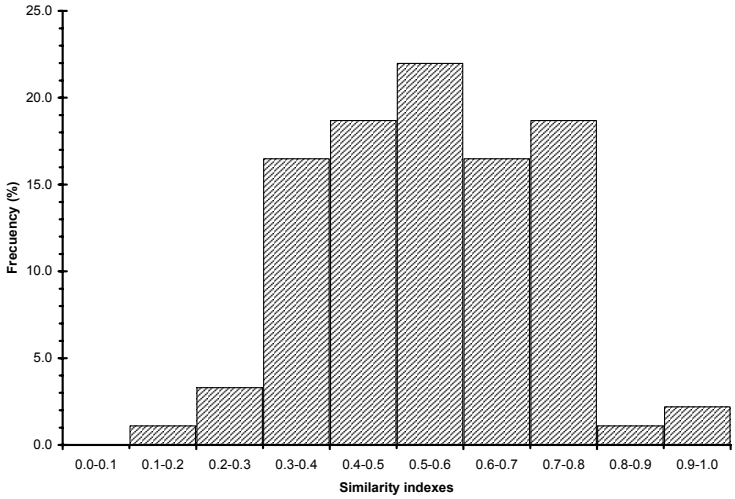
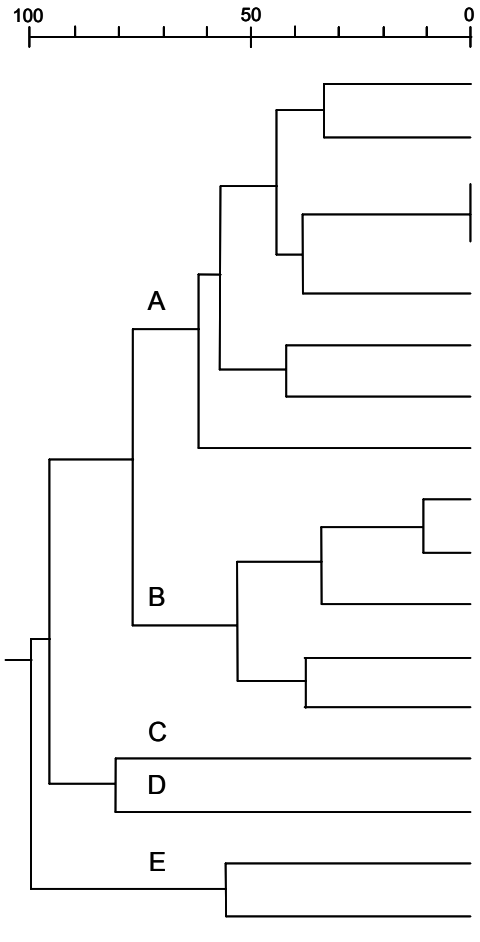


Fig. 3. Frequency histogram of the 91 similarity indexes grouped in 0.1 intervals.



strains	G418 (mg/ml)	TP (mg/l)	%	RP (mg/l)
GS115	0.00	--	--	--
pPIC9Kb	0.25	--	--	--
C1-5.5	0.25	34.0	21.1	7.2
C1-2.4	1.50	29.6	45.3	13.4
C1-3.5	1.00	n.d.	n.d.	n.d.
C2-4.1	0.75	34.2	51.6	17.6
pPIC9Ka	0.25	--	--	--
C2-6.2	0.00	n.d.	n.d.	n.d.
C2-5.3	1.00	27.4	29.2	8.0
C2-1.1	1.75	17.6	42.9	7.5
C1-6.1	4.00	21.6	23.8	5.1
C2-3.1	1.50	26.9	23.1	6.2
C3-6.1	1.50	n.d.	n.d.	n.d.
C1-1.3	0.25	n.d.	n.d.	n.d.
C2-5.4	0.75	31.0	23.3	7.2
C1-6.4	0.25	n.d.	n.d.	n.d.
C1-1.2	3.00	22.4	51.1	11.4

Fig. 4. Dendrogram plot showing strain clusters according to the band patterns. Cluster A: strains C1-5-5, C1-2.4, C1-3.5, C2-4.1 and the three control strains. Cluster B: strains C2-5.3, C2-1.1, C1-6.1, C2-3.1 and C3-6.1. Cluster C: strain C1-6.4. Cluster D: strain C2-5.5. Cluster E: strains C1-6.4 and C1-1.2. First column: strain identification. Second column: maximum G418 concentration (mg/ml) to which each strain was resistant. Third column: total protein (TP) concentration (mg/l). Fourth column: relative abundance of the recombinant protein (RP), with respect to the TP content. Fifth column: RP levels (mg/l). The last three data columns were calculated from the cell-free culture medium after 140-h shake-flask cultures in BMM and each value represents the mean of three results (coefficient of variation less than 5%) from three independent induction experiments. n.d.: not determined.

was the most frequently seen (17 of 17, 100%), followed by a 467-bp band, 15 times (88%); a 445-bp band, 14 times (82%); a 595-bp band, 13 times (76%); a 830-bp band, 12 times (71%); a 393-bp band, 11 times (65%); and a 292-bp band, 9 times (53%). Six of the seven more frequently seen bands were also present in the band pattern of the GS115 strain. Although all the recombinant

strains came from the same transformation assay of *P. pastoris* strain GS115 with the same vector, only two of the 14 (14%) recombinant strains had the same band pattern. Therefore, almost all of the strains (86%) were genotyped differently, and different from the control strains.

A frequency histogram of the 91 similarity indexes grouped in 0.1 intervals is shown in Fig. 3. Only three similarity indexes showed values greater than 0.80, while the majority of them (84 of 91, 92.3%) had values from 0.30 to 0.80, indicating a clear difference from the strains that generated these patterns, even though these strains still shared many bands.

Fig. 4 shows the dendrogram plot of the analyzed strains, where five clusters (A to E) were defined, grouping strains with a percentage of distance of less than 65% along the dendrogram scale. The strains belonging to cluster A are the strains most similar to the three control strains, and this group shows the greatest recombinant protein production (13.4 and 17.6 mg/l). The recombinant protein levels of strains of cluster B ranged from 5.1 to 8.0 mg/l. Although two strains of cluster A (C1-5.5 and C1-2.4) were identical, according to the AFLP analysis, these strains showed different resistances to G418, and different levels of recombinant protein production, indicating that our simplified AFLP method did not detect all of the genetic differences between the strains. No correlation was seen between resistance to G418 and the recombinant protein levels ($r^2 = 0.06$).

DISCUSSION

Genotyping techniques generate DNA fingerprints by creating DNA fragment patterns that can be separated by electrophoresis. The differences in the nucleotide sequences of the DNA are observed as the absence or presence of bands. AFLP is a powerful genotyping technique based on the detection of genomic restriction fragments amplified by PCR of any source or complexity of DNA. After the DNA to be evaluated is digested with one or two restriction enzymes, the resulting fragments are amplified with primers designed to recognize the restriction fragment ends. To carry out this amplification, the restriction fragment ends are first modified by the addition of oligonucleotide adapters that allow the amplification of a subset of the DNA fragments. PCR primers hybridize to the adapter and restriction site sequences, and also have extra nucleotides at their 3' end. Due to the very high reproducibility of the AFLP method [11], which we tested and demonstrated with our AFLP method applied to the *P. pastoris* strain GS115, variations in banding patterns are known to be direct reflections of the genetic differences between the strains examined. Therefore, they can be considered genomic fingerprints and useful for a numerical analysis. The AFLP technique has been reported to be successful in genetically characterizing a great variety of organisms [11-15]. Following the principle of Vos *et al.* [11], the adapter used in this work was designed so that the original restriction site would not be restored after ligation of the adapter to a restriction fragment. Most authors that perform the AFLP technique use two

restriction enzymes, two or more adapters, several PCR primers, polyacrylamide gels, and detection by radioactivity or fluorescence [11]. The method developed here uses only one restriction enzyme, one double strand adapter, one PCR primer, agarose electrophoresis gels, and a simple detection method with ethidium bromide stain. Thus, this method is powerful with respect to *P. pastoris* discrimination strains, and can be performed at a lower cost than other methods currently used for genotyping.

In this study, the AFLP technique for genotyping *P. pastoris* strains was employed for the first time. The results demonstrate the possibility of differentiating several recombinant strains of *P. pastoris* through this technique, even strains obtained from the same transformation assay using a single expression vector. The finding of 16 band patterns from the 17 evaluated strains confirms the hypothesis that a clonal variation was obtained due to genetic differences generated during the transformation event. Nevertheless, these genotypic differences did not correlate with resistance to G418, and therefore also not with the copy number of the expression cassette integrated into the *P. pastoris* genome. In addition, recombinant protein levels of shake-flask cultures in BMM of the transformants did not correlate with resistance to G418 as was described for other recombinant *P. pastoris* strains [8, 22]. Our findings clearly indicate that random mutations are generated in the *P. pastoris* genome during the transformation process, likely non-controlled integration events which affect the genotype of the transformants as well as the production levels of the recombinant protein. Thus, the spheroplast method of transformation of *P. pastoris* is not a controlled procedure as has been suggested in the literature [17]. This observation is important since these events could also take place in the processes where DNA is introduced in a host. Furthermore, the UPGMA cluster analysis shows that the strains with lesser genetic differences with respect to the *P. pastoris* host strain GS115 (cluster A) are the recombinant strains with the greatest recombinant protein production. This is in contrast with the strains with more genetic differences with respect to the *P. pastoris* host strain GS115 (cluster B), which produced 2.2 to 3.5 times less recombinant protein than the best strain of cluster A.

The banding patterns observed here are probably due to the addition of the amplified products of several restriction fragments of similar sizes, so that the visualization of amplified products in polyacrylamide gels could generate even larger band numbers, increasing the discrimination power of recombinant strains. In addition, polyacrylamide gels could be used to identify the genomic DNA region of the host where it has been altered during the *P. pastoris* transformation through the molecular cloning and sequencing of specific bands [23]. Thus, a greater understanding may be possible of the kind of uncontrolled events that occur during host transformation.

Acknowledgments. This study was supported by grant CN528-01 (PAICYT) from the Universidad Autónoma de Nuevo León. We are grateful to Carmen

Amelia Molina-Torres for technical support and Glen D. Wheeler for his stylistic suggestions in the preparation of this manuscript. We would like to dedicate this paper to the memory of Prof. R.M. Chandler-Burns who, for ten years, helped us with the English proofreading of our scientific papers.

REFERENCES

1. Romanos, M.A. Advances in the use of *Pichia pastoris* for high level gene expression. **Curr. Opin. Biotechnol.** 6 (1995) 527-533.
2. Escamilla-Treviño, L.L., Viader-Salvadó, J.M. and Guerrero-Olazarán M. Producción de proteínas recombinantes en *Pichia pastoris*. **CIENCIA UANL** 2 (1999) 27-33.
3. Cregg, J.M., Barringer, K.J., Hessler, A.Y. and Madden, K.R. *Pichia pastoris* as a host for transformations. **Mol. Cell. Biol.** 5 (1985) 3376-3385.
4. Clare, J.J., Rayment, F.B., Ballantine, S.P., Sreekrishna, K. and Romanos M.A. High-level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. **Biotechnology** 9 (1991) 455-60.
5. Clare, J.J., Sreekrishna, K. and Romanos M. Expression of tetanus toxin fragment C. **Methods Mol. Biol.** 103 (1998) 193-208.
6. Brankamp, R.G., Sreekrishna, K., Smith, P.L., Blankenship, D.T. and Cardin, A.D. Expression of a synthetic gene encoding the anticoagulant-antimetastatic proteinghilanten by the methylotropic yeast *Pichia pastoris.* **Protein Expr. Purif.** 6 (1995) 813-820.
7. Romanos, M., Scorer, C., Sreekrishna, K. and Clare, J. The generation of multicopy recombinant strains. in: **Pichia Protocols. Methods in Molecular Biology.** (Higgins, D.R. and Cregg, J.M., Eds.) Vol. 103, Humana Press Inc, Totowa, NJ, 1998, 55-72.
8. Sreekrishna, K., Brankamp, R.G., Kropp, K.E., Blankenship, D.L., Tsay, J.-T., Smith, P.L., Wierschke, J.D., Subramaniam, A. and Birkenberger, L.A. Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. **Gene** 190 (1997) 55-62.
9. Scorer, C.A., Buckholz, R.G., Clare, J.J. and Romanos, M.A. The intracellular production and secretion of HIV-1 envelope protein in the methylotrophic yeast *Pichia pastoris.* **Gene** 136 (1993) 111-119.
10. Scorer, C.A., Clare J.J., Macombie, R., Romanos, M.A. and Sreekrishna, K. Rapid selection using G418 of high copy number tranformats of *Pichia pastoris* for high-level foreign gene expression. **Biotechnology** 12 (1994) 181-184.
11. Vos, P., Hogers, R., Bleeker, M., Reijans, V., Van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. AFLP: a new technique for DNA fingerprinting. **Nucleic Acids Res.** 23 (1995) 4407-4414.

12. Lin, J.J. and Kuo, J. AFLP: a novel PCR-based assay for plant and bacterial DNA fingerprinting. **Focus** 17 (1995) 66-70.
13. Becker, J., Vos, P., Kuiper, M., Salamini, F. and Heun, V. Combined mapping of AFLP and RFLP markers in barley. **Mol. Gen. Genet.** 249 (1995) 65-73.
14. Janssen, P., Coopman, R., Huys, G., Swings, J., Bleeker, M., Vos, P., Zabeau, M. and Kersters, K. Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. **Microbiol.** 142 (1996) 1881-1893.
15. Ganter, P.F. and de Barros-Lopes, M. The use of anonymous DNA markers in assessing worldwide relatedness in the yeast species *Pichia kluyveri* Bedford and Kudrjavzev. **Can. J. Microbiol.** 46 (2000) 967-980.
16. Anonymous. *Pichia* expression kit: a manual of methods for expression of recombinant proteins in *Pichia pastoris*. Invitrogen**.** Available from: URL: http://www.invitrogen.com/content/sfs/manuals/pich_man.pdf.
17. Cregg, J.M. and Russell, K.A. Transformation. in: **Pichia Protocols. Methods in Molecular Biology.** (Higgins, D.R. and Cregg, J.M., Eds.) Vol. 103, Humana Press Inc, Totowa, NJ, 1998, 27-39.
18. Escamilla-Treviño, L.L., Viader-Salvadó J.M., Barrera-Saldaña H.A. and Guerrero-Olazarán, M. Biosynthesis and secretion of recombinant human growth hormone in *Pichia pastoris*. **Biotechnol. Lett.** 22 (2000) 109-114.
19. Hecker, K.H. and Roux, K.H. High and low annealing temperatures increase both specificity and yield in touchdown and stepdown PCR. **BioTechiniques** 20 (1996) 478-485.
20. Felsenstein, J. PHYLIP Phylogeny Inference Package (Version 3.2). **Cladistics** 5 (1989) 164-166.
21. Sneath, P.H.A. and Sokal, R.R. **Numerical Taxonomy: the principles and practice of numerical classification.** W.H. Freeman and Company, San Francisco, 1973, 230-234.
22. Sreekrishna, K. Strategies for optimizing protein expression and secretion in the methylotrophic yeast *Pichia pastoris*. in: **Industrial microorganisms: basic and applied molecular genetics.** (Balz, R.H., Hegeman, G.D. and Skatrud, P.L., Eds.) American Society for Microbiology, Washington, D.C., 1993, 119-126.
23. Knorr, C., Cheng, H.H., and Dodgson, J.B. DNA cloning and sequence analysis of chicken AFLP. **Anim. Genet.** 32 (2001) 156-159.