Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1014 (2003) 29–35
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Differential gene expression analysis by micro-preparative capillary gel electrophoresis

Andras Guttman*, Liang Shi, Julia Khandurina, Xun Wang
Torrey Mesa Research Institute, 3115 Merryfield Row, San Diego, CA 92121, USA

Abstract

Differential display analysis by cDNA fractionation, collection of differentially expressed fractions of interests and their downstream characterization is demonstrated. cDNA pools from two strains of *Cochliobolus heterostrophus* fungus were generated by specific restriction digestion and selective ligation. Micropreparative separation and isolation of differentially expressed transcript representatives were accomplished by high-performance capillary gel electrophoresis. The collected individual DNA molecules were polymerase chain reaction amplified and sequenced to create expressed sequence tags for the genes of interests. High resolving power and sensitivity of capillary gel electrophoresis enabled fast and automated processing of minute amounts of cDNA samples with high precision.

Keywords: Genes; *Cochliobolus heterostrophus*; DNA

1. Introduction

Capillary gel electrophoresis (CGE) was originally developed as a separation tool for biopolymer analysis, however, recent reports emphasized applicability of this method to micropreparations, especially in genomics and transcriptomics applications [1–3]. The high performance of gel or polymer solution filled narrow bore channels enable separations of DNA molecules of very similar size with sequencing grade, i.e., single base resolution up to several hundred bases in fragment size. Moreover, micropreparative CGE readily provides enough material for downstream sample processing by polymerase chain reaction (PCR), cycle sequencing and cloning [4]. Recent developments in capillary electrophoresis instrumentation made possible collection of closely migrating peaks with enhanced precision using microfluidic adapters and their combinations with long separation capillaries [5].

The simplest fraction collection approach in capillary gel electrophoresis employs constant electric field during both the separation and fractionation processes. Collection timing and duration are defined by the migration velocity of the detected peak of interest, column length, and analyte zone width in time units, respectively. To accommodate the high peak capacity of CGE, voltage programming was introduced to improve fraction collection precision by maintaining high electric field for the separation but slowing down the electrophoretic migration during collection [6]. Sheath flow support was attempted to better assist the fractionation process [7] especially in conjunction with gel filled capillary columns where there is no apparent fluid flow at the

*Corresponding author. Present address: Diversa, 4955 Directors Place, San Diego, CA 92121, USA. Tel.: +1-858-526-5421; fax: +1-858-526-5921.
E-mail address: aguttman@diversa.com (A. Guttman).

0021-9673/03/$ – see front matter
doi:10.1016/S0021-9673(03)00777-5

column outlet. Later, other approaches were introduced to increase collection precision, such as double point detection [8]. A simple time interval based collection also proved to be a feasible approach. For large-scale applications, 96-well plate format was used with voltage interruption for time-based microfractionation followed by consequent downstream fraction processing and characterization [9]. Please note that this approach did not require detection during fractionation but the collected samples were subject to PCR amplification followed by either high throughput agarose gel electrophoresis or fluorescent staining in conjunction with a microplate reader to identify the wells containing amplified fragments. To further increase fractionation capabilities, a multicapillary system was introduced by Minarik et al. to accommodate collection of hundreds to thousands of DNA fragments in predetermined time intervals at various stages of the separation [10]. To prevail interruption of the electric field between the separation and collection cycles in large-scale collection of multiple DNA fragments, they introduced agarose microwell plates that enabled continuous fractionation without high-voltage interruption.

Collection of multiple DNA fragments is very important during gel electrophoresis based expressed sequence tag (EST) library generation processes. Apparent problems with conventional EST methodologies are redundancy of highly expressed transcripts, difficulty to monitor low abundance transcripts, and the high cost associated with mapping non-model organisms. cDNA microarrays and genechip technology, widely used for EST generation, however, both require apriori sequence information [11]. Serial analysis of gene expression (SAGE) is not based on sequence knowledge, but generates only short sequence tags, thus, necessitating extensive follow-up verification steps [12]. The methods of choice for studying genomes with unknown sequences are amplified fragment length polymorphism (AFLP) [13] and differential display analysis [14]. While none of these techniques require apriori sequence information, high-resolution separation and fractionation capabilities are needed to establish libraries. Traditional approach employs time consuming slab gel electrophoresis followed by tedious manual band excision and DNA extraction. Novel methodologies take advantage of capillary electrophoresis with automated fraction collection capabilities in 96-well plate format eliminating excision and extraction steps.

Southern corn leaf blight disease is caused by the fungus *Cochliobolus heterostrophus*, that has two nearly isogenic strains of Tox^- and Tox^+ [15]. The high virulence of Tox^+ strain to Texas male sterile corn (T-corn) is caused by the presence of *Tox1* locus generating T-toxin [16]. Since the Tox^- strain lacks the *Tox1* locus, it does not produce the toxin, therefore, exhibits only weak virulence. Polyketide synthase (*PKS1*) and decarboxylase (*DEC1*) [17] are among the most important genes directly associated with T-toxin production. Another gene, reductase (*RED1*), was also found to be tightly linked to *DEC1*, but not necessary for toxin synthesis. All three genes are present in the Tox^+ strain but absent in the Tox^- strain of *C. heterostrophus*.

In this paper we present a methodology for differential gene expression analysis using micropreparative capillary gel electrophoresis. Transcription profiles of the Tox^- and Tox^+ strains of *C. heterostrophus* were generated and analyzed. The differentially expressed gene transcripts were then collected, processed by PCR amplification and sequenced.

2. Materials and methods

2.1. Sample preparation

Total RNA was extracted from two near isogenic strains, Tox^- and Tox^+, of the plant pathogenic fungus *C. heterostrophus* according to our earlier published procedure [17]. RNA quality and concentration were assessed by capillary electrophoresis [18].

2.2. Ligation specificity based differential gene expression analysis

Double-stranded cDNA was first synthesized from *C. heterostrophus* total RNA, using the Superscript cDNA synthesis kit (Life Technologies), followed by digestion with *Bsa*J I (C/CNNGG) in a combination with a six-base cutter *Eco*R I (G/AATTC) restriction enzymes, resulting in four-base sticky overhangs. Both enzymes were from New England Biolabs

(Beverly, MA, USA). Each reaction mixture contained 40 ng cDNA, 1 unit *Eco*R I and 4 units *Bsa*J I, 1× restriction buffer with bovine serum albumin (BSA), in a final volume of 34 μl. The reaction was incubated at 37 °C for 1.5 h, followed by an additional 1.5 h at 60 °C. The digested double-stranded cDNA fragments were then ligated to double stranded specially designed adapters. To accommodate DNA fragment overhang ends, formed by *Eco*R cleavage, the upper strand of the adapter was designed to prevent restoring the restriction enzyme recognition site after the ligation (5′-GCTGCTAGTGTCCGATGT/A-3′). The lower strand comprised a common sequence, complementary to the upper strand, and the enzyme specific restriction site sequence. For the ligation of overhangs generated by *Bsa*J I digestion, the upper adapter sequence was 5′-GATCTCCTAGAGTCGTGA-3′, while the lower one was complementary to the upper strand, plus the restriction site sequence. Due to the two-base degeneracy in the recognition sequence, 16 different ends ($N^2=16$) were generated by *Bsa*J I digestion, i.e., there were 16 different ligation reactions (sub-pools) for each double-digested cDNA pool. In each ligation sub-pool, the six-base cutter enzyme adapter was paired with one of the 16 different adapters, matching perfectly one of the possible *Bsa*J I subset combinations. Each reaction contained 1–2.5 ng digested cDNA, 0.1 pmol *Eco*R I enzyme adapter, 0.1 pmol of *Bsa*J I adapter (one of the sixteen kinds), 1.5 Weiss units T4 DNA ligase (New England Biolabs) and 1× ligation buffer, in a final volume of 8 μl. Following the addition of T4 DNA ligase, the reaction mixtures were incubated at 16 °C for 2 h. The ligated fragments were then PCR amplified with a set of common primers containing sequences identical to the upper strand adapters, with one of the primers labeled at the 5′ end with fluorescein (Applied Biosystems, Foster City, CA, USA). Each PCR contained 1–2.5 ng of adapter-ligated cDNA, 1× PCR buffer, 1.5 m*M* $MgCl_2$, 200 μ*M* deoxyribose nucleotide triphosphate (dNTP), 6% dimethylsulfoxide (DMSO), 20 pmol of each primer, and 2.5 units Taq DNA polymerase (Life Technologies), in a total volume of 25 μl. Thermocycling program started with a 5 min denaturation step at 94 °C, followed by 25 cycles of 94 °C for 30 s, 60 °C for 30 s and 72 °C for 60 s.

2.3. Capillary gel electrophoresis separation and fraction collection

Amplified PCR products (cDNA fingerprints) were size separated using a P/ACE MDQ system (Beckman Coulter, Fullerton, CA, USA) under reversed polarity settings (anode at the detection side and cathode at the injection end). Separations were monitored on-column by laser-induced fluorescence (LIF) with an Ar-ion laser at 488 nm excitation and 520 nm emission wavelengths. The data were acquired and processed by the P/ACE MDQ 32 Karat software package (Beckman Coulter). A bare fused-silica capillary (75 μm I.D., Polymicro Technologies, Phoenix, AZ, USA) was used, with the total length (from the injection to the collection vial) of 60 cm (50 cm to the detection window). The sieving polymer POP-5 was introduced into the capillary by pressure (80 p.s.i.; 1 p.s.i.=6894.76 Pa) and replaced after each run. The capillary cartridge was thermostated at 40.0±0.1 °C by active liquid cooling. The applied electric field during the separation and collection steps was 175 V/cm. Fragments were collected based on their migration velocity and the calculated time as they are supposed to reach the end of the column [4]. The fractions were collected into individual 200-μl microfuge tubes containing 5 μl 0.1× TBE buffer. Tris, boric acid and EDTA for the 1× TBE buffer (89 m*M* Tris, 89 m*M* boric acid, 2 m*M* EDTA·Na_2, pH 8.4) and formamide were from ICN (Costa Mesa, CA, USA). A Genescan 500 ROX fluorophore labeled sizing DNA ladder, POP-5 sequencing sieving matrix and run buffer were from Applied Biosystems. All buffer solutions were filtered through a 0.2-μm filter before use (Schleicher and Schuell, Keene, NH, USA). Samples were dissolved in 85% formamide and denatured at 95 °C for 5 min prior to electrokinetic injection at 10 kV for 10–60 s. Re-injection analysis of the collected and PCR amplified DNA fragments was carried out under the same settings.

2.4. Amplification and sequencing of the collected fragments

For the amplification of the collected differentially expressed DNA fragments, 47 μl PCR master mix, containing 45 μl Platinum PCR supermix (Invitrogen, Carlsbad, CA, USA) and 1 μl of each PCR

primer at 10 μ*M* concentration, was added to 3 μl of each fraction. The sequences of the PCR primers were as follows:

5′ - GAAACAGCTATGACCATGGCTGCTAGT-GTCCGATGT-3′ and 5′-TGTAAAACGACGGCC-AGTGATCTCCTAGAGTCGTGA-3′.

Thermal cycling conditions were the same as in Section 2.1. PCR products were first analyzed by CGE (conditions were the same as in Section 2.2) and then sequenced on an ABI 3700 instrument (Applied Biosystems) using M13 forward and reverse primers [17].

3. Results and discussion

The open architecture ligation specificity based expression profiling technique used for differential display analysis is schematically outlined in Fig. 1. Please note that this method did not require any apriori sequence information about the species of interest. First, a cDNA library was generated from the relevant mRNA pools of the *Tox*$^{+}$ and *Tox*$^{-}$ strains of *C. heterostrophus*. cDNA was then fragmented with a carefully selected restriction enzyme pair of *Eco*R I and *Bsa*J I and ligated with specific adapters, reducing in this way the typical cDNA population by two to three orders of magnitude, to the complexity level suitable for CGE-based analysis. This was followed by amplification of the ligated cDNA fragments using fluorescein labeled primers. High-resolution separation and detection of the resulted DNA fragments was accomplished by automated capillary gel electrophoresis. Fig. 2 exhibits CGE traces of a typical ligation specificity based expression analysis sample (lower trace), and a DNA sizing standard (upper trace) for accurate fragment length assignment.

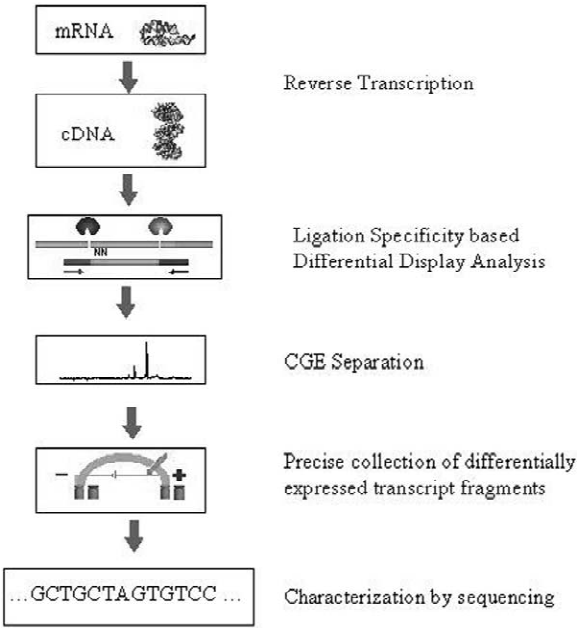

Fig. 1. Schematics of the open architecture ligation specificity based differential gene expression profiling technique.

To identify differentially expressed genes, transcription profiles of the two strains of *C. heterostrophus* were generated by ligation specificity based expression analysis. Fig. 3 compares a selected portion of the two profiles of *Tox*$^{-}$ (upper trace) and *Tox*$^{+}$ (lower trace) strain samples. Quantitative signal intensity comparison was accomplished automatically using the data acquisition and processing capabilities of the CE instrument, i.e., fluorescent signal was normalized to the maximum signal intensity among the samples. Height and area of each individual peaks, representing different gene expression levels were compared to controls and peaks with more than twofold differences were considered as differentially expressed. As one can observe, the upper trace in Fig. 3 shows only one major peak (F2), while the lower trace has two major fragments (F1 and F2), suggesting differential expression of the genes associated with these transcripts, between the two species studied. Both F1 and F2 fragments were collected by the above described micropreparative CGE procedure for re-amplification and sequencing.

Fraction collection was accomplished, based on the measured migration times of the peaks of interest at the detection point. When the onset of the analyte zone of interest was expected to reach the capillary outlet, the capillary end was transferred from the buffer reservoir to the corresponding collection vial. The precise timing of the vial switch was calculated based on the migration velocity of the peak to be collected and the total to effective (to the detection point) capillary length ratio. The duration of the collection step corresponded to the band-width (in time units) of the sample zone of interest. Collection

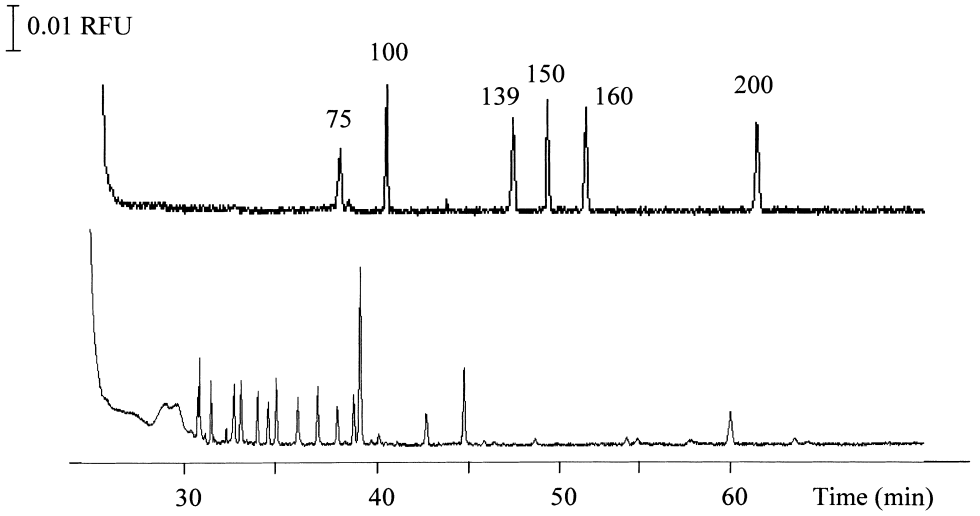

Fig. 2. Comparison of the capillary gel electrophoresis profiles of ROX-labeled sizing ladder (upper trace) to a fluorescein-labeled ligation specificity based differential gene expression analysis sample (lower trace). Capillary: bare fused-silica, 60 cm (50 cm from injection to detection)×75 μm I.D., LIF detection: 488 nm excitation/520 nm emission. Denaturing sieving matrix: POP-5. Separation temperature: 40 °C. Samples were prepared in 85% formamide (95 °C/5 min). Electrokinetic injection: 10 kV for 10 s. Separation voltage: 10.5 kV (reversed polarity). Numbers above the peaks correspond to base numbers.

steps were performed under the same applied voltage as used for separation. Fig. 4 demonstrates the separation, fraction collection, re-injection and purity analysis processes. Trace A in Fig. 4 depicts the corresponding electropherogram section of the ligation specificity based differential expression analysis sample, containing the two fragments of interest. Fractions were collected for the time intervals shown above trace A, i.e., 35 and 45 s for F1 and F2, respectively. Verification of the successful collection process was accomplished by re-injection of the fractions. Traces B and C show the electropherograms of the re-injected fractions, demonstrating the purity of the collected DNA fragments showing no cross-contamination. Please note that since the amount of the collected and re-injected DNA by CGE was miniscule, for better visualization, the signal of traces B and C was scaled up 20-fold. The collected samples were then PCR amplified to confirm feasibility of downstream sample processing, after micropreparative CE fractionation. Since all ligated DNA fragments possessed the same flanking sequences, the very same primer set was used in all PCRs. Traces D and E show the electropherograms of the amplified products, matching well to F1 and F2 fragment sizes. A blank PCR sample was also injected as a negative amplification control (trace F). The collected and amplified fractions (F1 and F2) were sequenced on an ABI 3700 system, and were

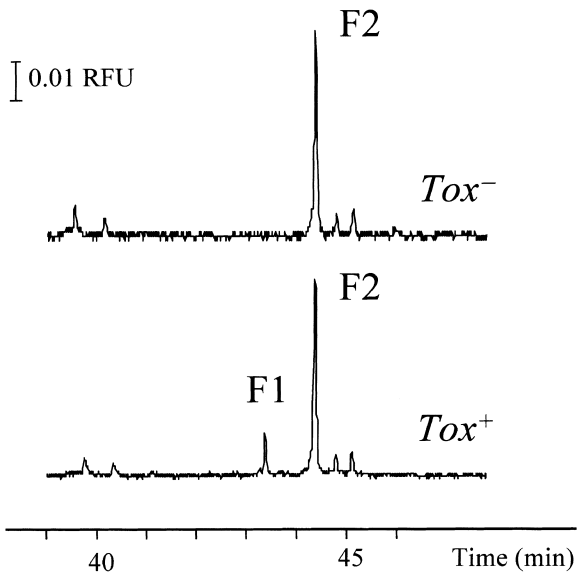

Fig. 3. Differential display analysis of the *Tox*$^{-}$ and *Tox*$^{+}$ strains of *Cochliobolus heterostrophus* fungus by CGE. Separation conditions as in Fig. 2.

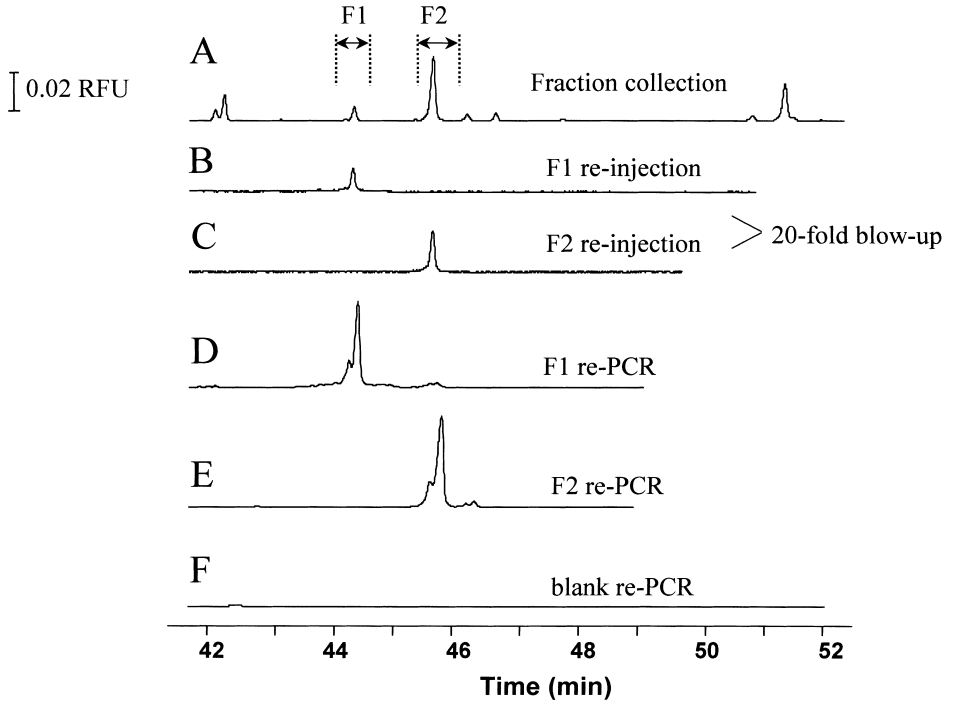

Fig. 4. Fraction collection and purity verification by capillary gel electrophoresis. (A) Electropherogram section of a ligation specificity based differential gene expression analysis sample containing two peaks of interest (F1 and F2). Collection times: 35 s for F1 and 45 s for F2. (B, C) Re-injection of the collected fractions F1 and F2, respectively (signal scaled up 20-fold). (D, E) PCR amplification products of the collected fractions F1 and F2, respectively. (F) Blank PCR (no template) sample as a negative control. Separation conditions as in Fig. 2.

apparently long enough (114 and 122 base pairs, respectively) to be used as expressed sequence tags for bioinformatics based homology searches.

4. Conclusions

Micropreparative capillary gel electrophoresis has been applied to collect and sequence differentially expressed gene transcript fragments, providing a method for monitoring gene expression levels from samples of different environmental, developmental and physiological conditions. cDNA pools were generated by an open architecture sub-pooling methodology, based on fragmentation of cDNA populations by special restriction enzyme combination to create different DNA fragment overhangs, followed by specific ligation and amplification. By this means the level of complexity of cDNA fragment mixtures was substantially reduced, enabling separation and collection of individual differentially expressed transcript representatives at microscale by CGE, and their further downstream processing by PCR and characterization by sequencing. The method introduced here overcame the limitation of microarray techniques in distinguishing between the variation of species of a gene family, as well as the cross-hybridization by gene family members, thus provided flexible differential gene expression method. Since each fragment in the profile had a defined length and sequence, it could be readily linked to other genome databases electronically.

References

[1] P.O. Ekstrom, R. Wasserkort, M. Minarik, F. Foret, W.G. Thilly, Biotechniques 29 (2000) 582.
[2] T. Irie, T. Oshida, H. Hasegawa, Y. Matsuoka, T. Li, Y. Oya, T. Tanaka, G. Tsujimoto, H. Kambara, Electrophoresis 21 (2000) 367.

[3] S. Magnusdottir, C. Heller, P. Sergot, J.L. Viovy, Electrophoresis 18 (1997) 1990.
[4] Z. Ronai, A. Guttman, Am. Lab. 32 (2000) 28.
[5] J. Khandurina, A. Guttman, J. Chromatogr. A 979 (2002) 105.
[6] A. Guttman, A.S. Cohen, D.N. Heiger, B.L. Karger, Anal. Chem. 62 (1990) 137.
[7] C. Schwer, F. Lottspeich, J. Chromatogr. 623 (1992) 345.
[8] M. Minarik, F. Foret, B.L. Karger, Electrophoresis 21 (2000) 247.
[9] L. Shi, J. Khandurina, Z. Ronai, B.-Y. Li, W.K. Kwan, X. Wang, A. Guttman, Electrophoresis 24 (2003) 86.
[10] M. Minarik, K. Kleparnik, M. Gilar, F. Foret, A.W. Miller, Z. Sosic, B.L. Karger, Electrophoresis 23 (2002) 35.
[11] D.J. Lockhart, Nat. Med. 4 (1998) 1235.
[12] V. Velculescu, L. Zhang, B. Vogelstein, Science 270 (1995) 484.
[13] P. Breyne, M. Zabeau, Curr. Opin. Plant Biol. 4 (2001) 136.
[14] P. Liang, A.B. Pardee, Science 257 (1992) 967.
[15] M. Kodama, M.S. Rose, G. Yang, Genetics 151 (1999) 585.
[16] B.G. Turgeon, S.W. Lu, in: J.W. Kronstad (Ed.), Fungal Pathology, Kluver, New York, 2000, p. 93.
[17] B.-Y. Li, W.K. Kwan, B.G. Turgeon, J. Wu, X. Wang, E. Li, T. Zhu, L. Shi, Omics 6 (2002) 175.
[18] J. Khandurina, H.-S. Chang, B. Wanders, A. Guttman, Biotechniques 32 (2002) 1226.