Review article

An Integrated Approach to Functional Genomics: Construction of a Novel Reporter Genomic Library for *Osmium baselicum* Plant

Gulab Chand Shah*[1], Vikrant Gupta[2], SPS Khanuja[2]

*corresponding author- E-Mail: gulab777@gmail.com

1 Rajiv Gandhi College Satna M.P. India, 2 Central Institute of Medicinal and Aromatic Plant Lucknow U.P. India

_____

Abstract

As a means of investigating gene function, we developed a vigorous transcription fusion reporter vector to measure gene expression in plant. The vector, plasmid, was used to construct a haphazard insert library for the *Osmium bsilicum* genome. plasmid replicates in *Escherichia coli* and can be transferred to, but cannot replicate in, *S. meliloti*. Homologous recombination of the DNA fragments cloned in plasmid into the *Osmium bsilicum* genome generates transcriptional fusions to either the reporter genes *gfp* and *lacZ* or *gusA* and *rfp*, depending on the orientation of the cloned section. A database containing all the gene expression activities together with a web interface showing the precise locations of reporter fusion junctions has been constructed. Sequence study, and the plasmid clones were recombined into *Osmium bsilicum*. Reporter enzyme activities following growth of these recombinants in complex medium (LB) and in minimal medium with glucose or succinate as the sole carbon source allowed the identification of genes highly expressed under one or more growth condition and those uttered at very low to background levels. In addition to generating reporter gene fusions, the vector allows Flp recombinase-directed deletion formation and gene disruption, depending on the nature of the cloned fragment. We report the identification of genes essential for growth on complex medium as deduced from an in capacity to recover recombinants from plasmid clones that carried fragments internal to gene or operon transcripts.

Keywords: Genome, Library, Ammonia, Promoter, Transcriptome
_____

Introduction

Microbial genome sequences have exposed a massive amount of genes for which purposeful information is almost wholly lacking. One can infer that the functions of these genes are assorted, and hence, at a whole-genome level, it would seem sensible to adopt multiple approaches to the identification of gene functions. Several global approaches

1

are being applied to the analysis of gene function in model organisms. These consist of in silico bioinformatics studies (2, 5), systematic gene interruption (3), transcriptome microarray studies (1, 4, 6, 8), gene fusion and expression studies (3, 7), promoter shutoff (5), the application of protein fusion tags, proteomic analysis (6), and two-hybrid screens to identify interacting proteins (3). Genome-level studies of the legume symbiont *Osmium bsilicum* have also been under way. This organism induces nitrogen-fixing root nodules on alfalfa. During the interaction between the eukaryotic host and prokaryotic bacterium, signals are exchanged which harmonize gene expression in both organisms to give rise to the N-fixing nodule. The bacteria within the nodule are bounded by a host-derived cell membrane and are supplied with a reduced carbon energy source(s) from the plant, and in turn they bring in the plant with reduced nitrogen (7). As a free-living organism, the bacterium grows and survives in the soil surroundings and in the rhizosphere surrounding the plant root. In the laboratory, the free-living bacterium can grow in complex medium and in defined smallest media with glucose, succinate, or a broad range of carbon sources and can make the most of ammonia, nitrate, or a assortment of other compounds as sources of nitrogen.

The recent innovation and categorization of several large dsDNA viruses from aquatic environments belonging to the phycodnavirus and mimivirus families led to the expansion of new hypotheses (9-11). A comparative analysis of the gene contented of these viruses with poxviruses, iridoviruses and asfarviruses indicated that they have nine gene yield in common, and 33 more gene yield are present in at least two of these families. It follows that they might have a common evolutionary antecedent, a nucleocytoplasmic large dsDNA virus (NCLDV) (12, 13).

Bacterial strains and plasmids (14)

E. coli strains were grown on LB agar plates, supplemented with ampicillin (100 μg/mL) where necessary.

Fungal strains and growth conditions (14, 15)

Cultures of *Neotyphodium lolii* strain Lp19 (16) and *E. festucae* strain Fl1 (ex cultivar SR3000) were grown and maintained as described (17, 18).

Molecular biology techniques

Fungal genomic DNA was isolated from freeze-dried mycelia using previously described methods (19, 20). Plasmid DNA was isolated and purified by alkaline lysis using either the Bio-Rad (Hercules, CA 94547, USA) Quantum plasmid miniprep or midiprep kits or the Roche (Roche Diagnostics N.Z., Ltd., Auckland, New Zealand) plasmid Miniprep kit. Genomic DNA digests were transferred to positively charged nylo membranes (Roche) by capillary transfer (21) and fixed by UV cross-linking (120,000μJ/cm) in an Ultraviolet cross-linker Cex-800 (Ultralum, Inc., Claremont, CA 91711, USA). Filters were probed with [a 2 P]-dCTP (3000 Ci/ mmol; GE Healthcare, Auckland, New Zealand) labeled probes (Additional File 1).DNA was labeled by primed synthesis with Klenow fragment using a High-Prime kit (Roche). Labeled probes were purified using ProbeQuant™ columns (GE Healthcare). Membranes were washed and hybridization signals detected by autoradiography as described previously (22).

Gene cloning strategy

Nine SLP genes were identified in *E. festucae* Fl1 either using sequences amplified from the closely related fungal species *N. lolii* Lp19 as described previously (23), or amplified from *E. festucae* Fl1 genomic DNA with degenerate primers (Additional Files 1 and 2). Probes for these genes were hybridized to both an *E. festucae* Fl1 genomic DNA library and a Southern blot containing restriction enzyme digests of *E. festucae* Fl1 genomic DNA. Screening of the genomic library identified clones containing DNA of the gene of interest. Southern hybridizations provided information about the restriction enzyme fragmentscontaining the gene of interest, thus facilitating the subcloning of DNA fragments containing the desired gene. Six further SLP genes were identified in the genome of another *E. festucae* strain, E2368.

Library screening

building of the *N. lolii* Lp19 and *E. festucae* Fl1 genomic DNA libraries screened in this study was described previously [23,24]. The *N. lolii* Lp19 genomic DNA library was screened by plaque hybridization using typical methods (25). For the *E. festucae* Fl1 genomic library equipped as described in (24), filters arrayed with DNA from 5088 autonomous ampicillin-resistant colonies at a 6×6 density with double offset (Australian Genome Research Facility, Melbourne, Australia) were screened by hybridization with radioactively labeled probes (25).

Polymerase chain reaction and amplification conditions

Typical PCR amplifications of genomic DNA templates were carried out in 25 μL reactions containing 10 mM Tris-HCl, 1.5 mM MgCl2 and 50 mM KCl (pH 8.3), 50 μM of each dNTP, 200 nM of all primer, 0.5 U of *Taq* DNA polymerase (Roche) and 5 ng of genomic DNA. The thermocycler situation used were: 94°C for 2 min; 30 cycles of 94°C for 30 s, 60°C for 30 s and 72°C for 1 min per kb, followed by a concluding step at 72°C for 5 min.

DNA sequencing

DNA wreckage were sequenced by the dideoxynucleotide chain-termination method (26) using Big Dye (Version 3) chemistry with oligonucleotide primers (Sigma Genosys, Castle Hill, Australia) specific for pUC118, pGEM-T Easy, and genomic sequences from *N. lolii* or *E. festucae*. yield were separated on moreover an ABI Prism 377 sequencer (Perkin Elmer, Waltham, MA 02451, USA) or an ABI 3730 analyzer (Applied Biosystems, Inc., Foster City, CA 94404, USA) at the Allan Wilson Centre Genome Service, Massey University, Palmerston North, New Zealand.

Bioinformatic analyses

Sequence data were assembled into condign with SEQUENCHER (Gene Codes Corporation, Ann Arbor, MI 48108, USA) version 4.1 and analyzed and annotated using MacVector 7.2 (MacVector, Inc., Cary, NC 27519, USA). series comparisons were performed at the National Center of Biotechnology Information (NCBI) site http://www.ncbi.nlm.nih.gov using the Brookhaven (PDB), SWISSPROT, GenBank (CDS translation), PIR and PRF databases employing algorithms for both local (BLASTX and BLASTP) and worldwide (FASTA) alignments (27-29). Potential open appraisal frames for SLP and unlinked non-SLP genes were identified using FGENESH, an HMM-based gene structural guess using the *Fusarium graminearum* parameters http:linux1.softberry.coberry.phtml?topic=fgenesh&grp= ograms&sub group=gfind. There were some instances where SLPs were not annotated in genome sequences. TBlastN analysis, using a preserved region of the peptidase S8 domain as the query sequence, was used to identify all putative SLPs in genome sequences. Where supplementary SLP genes were acknowledged they were incorporated in the analysis. The attendance of signal peptides was analyzed using SignalP3.0 (30). Polypeptide alignments were performed by means of ClustalW (31) in MEGA4 (32). Phylogenetic analyses were conducted in MEGA4 (32). The evolutionary history was contingent using maximum likelihood (PhyML)(33). PhyML was run from the ATGC Montpellier Bioinformatics display place at http://www.atgcmontpellier.fr/. The Newick files were imported into MEGA 4.0 (32) to view the foliage which were saved in tif format. Sequence associations were inferred using the Neighbor-joining (N-J) method (34). The bootstrap N-J agreement tree inferred from 1000 replicates was taken to symbolize the evolutionary history of the taxa analyzed (35). Branches parallel to partitions reproduced in less than 50% bootstrap replicates were collapsed. The evolutionary distances were computed using the Poisson correction method (36) and are in the units of the number of amino acid substitutions per site. All positions containing alignment gaps and gone astray data were eliminated only in pairwise series comparisons (Pairwise deletion option). Assignment of subtilisin-like proteases to different families was done on the basis of sphere organization, similarity to other proteases and grouping in phylogenetic trees. Proteinase K type enzymes have propeptide and peptidase S8 domains. Subfamilies sf1, sf2 and sf3 were previously described (4). Subfamilies sf4 and sf5 of this group were assigned on the basis of their phylogenetic consortium. The pyrolysins have an S8 domain, intermittent by a PA domain, and a DUF1034 domain.

The two subfamilies within this cluster were assigned on the basis of previous work (4). The OSPs have a peptidase S8 sphere and distinct amino acid motifs exclusive to these relations (37).

Conclusion

In this study, we aimed to study the evolution of the SLP gene family in ocimum basilicum. Fifteen predicted SLP genes were present in the ocimum basilicum genome, representing four different SLP families. New subfamilies within the proteinase K family were identified, as well as a new family, the oxidatively stable proteases previously thought to be present only in bacteria. Phylogenetic studies showed that many gene duplications and loss events have occurred during evolution of the SLP gene family within the Hypocreales.

References

[1] Ampe, F., E. Kiss, F. Sabourdy, and J. Batut. 2003. Transcriptome analysis of Sinorhizobium meliloti during symbiosis. Genome Biol. 4:R15.

[2] Barnett, M. J., R. F. Fisher, T. Jones, C. Komp, A. P. Abola, F. Barloy- ubler, L. Bowser, D. Capela, F. Galibert, J. Gouzy, M. Gurjal, A. Hong, L. Huizar, R. W. Hyman, D. Kahn, M. L. Kahn, S. Kalman, D. H. Keating, C. Palm, M. C. Peck, R. Surzycki, D. H. Wells, K. C. Yeh, R. W. Davis, N. A. Federspiel, and S. R. Long. 2001. Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti pSymA megaplasmid. Proc. Natl. Acad. Sci. USA 98:9883–9888.

[3] Barnett, M. J., V. Oke, and S. R. Long. 2000. New genetic tools for use in the Rhizobiaceae and other bacteria. BioTechniques 29:240–245.

[4] Barnett, M. J., C. J. Toman, R. F. Fisher, and S. R. Long. 2004. A dualgenome symbiosis chip for coordinate study of signal exchange and development in a prokaryote-host interaction. Proc. Natl. Acad. Sci. USA 101: 16636–16641.

[5] Barsch, A., T. Patschkowski, and K. Niehaus. 2004. Comprehensive metabolite profiling of Sinorhizobium meliloti using gas chromatography-mass spectrometry.Funct. Integr. Genomics 4:219–230.

[6] Becker, A., H. Berges, E. Krol, C. Bruand, S. Ruberg, D. Capela, E. Lauber, E. Meilhoc, F. Ampe, F. J. de Bruijn, J. Fourment, A. Francez-Charlot, D. Kahn, H. Kuster, C. Liebe, A. Puhler, S. Weidner, and J. Batut. 2004. Global changes in gene expression in Sinorhizobium meliloti 1021 under microoxic and symbiotic conditions. Mol. Plant Microbe Interact. 17:292–303.

[7] Bjarnason, J., C. M. Southward, and M. G. Surette. 2003. Genomic profiling of iron-responsive genes in Salmonella enterica serovar Typhimurium by high-throughput screening of a random promoter library. J. Bacteriol. 185: 4973–4982.

[8] Blondelet-Rouault, M. H., J. Weiser, A. Lebrihi, P. Branny, and J. L. Pernodet. 1997. Antibiotic resistance gene cassettes derived from the omega interposon for use in E. coli and Streptomyces. Gene 190:315–317.

[9] Van, Etten J. L., Graves., M. V., Muller., D. G., Boland., W, Delaroque. N. Phycodnaviridae—large DNA algal viruses. Arch Virol 2002, 147(8):1479-1516.

[10] Raoult., D, Audic., S, Robert., C, Abergel., C, Renesto., P, Ogata., H, La., Scola B, Suzan., M, Claverie., J. M. The 1.2-megabase genome sequence of mimivirus. Science 2004, 306(5700):1344-1350.

[11] Wilson., W. H., Schroeder., D. C., Allen., M. J., Holden., M. T., Parkhill., J, Barrell., B. G., Churcher., C, Hamlin., N, Mungall., K, Norbertczak., H, Quail.,M. A., Price., C, Rabbinowitsch., E, Walker., D, Craigon., M,

Roy., D, Ghazal., P. Complete genome sequence and lytic phase transcription profile of a Coccolithovirus. Science 2005, 309(5737):1090-1092.

[12] Iyer, L. M, Aravind, L, Koonin, E. V. Common origin of four diverse families of large eukaryotic DNA viruses. J Virol 2001, 75(23):11720-11734.

[13] Iyer, L. M, Balaji, S, Koonin. E. V, Aravind, L. Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. Virus Res 2006, 117(1):156-184. Koonin EV, Senkevich TG, Dolja VV: The ancient Virus World.

[14] Michelle, K., Bryant, Christopher, L., Schardl, Uljana., Hesse and Barry Scott. Evolution of a subtilisin-like protease gene family in the grass endophytic fungus Epichloë festucae BMC Evolutionary Biology 2009, 9:168 doi:10.1186/1471-2148-9-168.

[15] Rawlings ND, Tolle DP, Barrett AJ: MEROPS: the peptidase database. Nucleic Acids Res 2004, 32:D160-164.

[16] Christensen MJ, Leuchtmann A, Rowan DD, Tapper BA: Taxonomy of Acremonium endophytes of tall fescue (Festuca arundinacea), Meadow fescue (F. pratensis) and perennial ryegrass (Lolium perenne). Mycol Res 1993, 97:1083-1092 .

[17] Moon, C. D, Tapper, B. A, Scott, B. Identification of Epichloë endophytes in planta by a microsatellite-based PCR fingerprinting assay with automated analysis. Appl Environ Microbiol 1999, 65:1268-1279.

[18] Moon CD, Scott B, Schardl CL, Christensen MJ. The evolutionary origins of Epichloë endophytes from annual ryegrasses. Mycologia 2000, 92:1103-1118.

[19] Byrd., A. D, Schardl., C. L, Songlin., P. J, Mogen., K. L, Siegel., M. R. The betatubulin gene of Epichloë typhina from perennial ryegrass (Lolium perenne). Curr Genet 1990, 18:347-354.

[20] Möller., E. M, Bahnweg., G, Sandermann., H, Geiger., H. H. A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. Nucleic Acids Res 1992, 20:6115-6116.

[21] Southern., E. M. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 1975, 98:503-517.

[22] Young, C., Itoh, Y., Johnson, R., Garthwaite, I., Miles, C. O, Munday-Finch S. C, Scott, B: Paxilline-negative mutants of Penicillium paxilli generated by heterologous and homologous plasmid integration. Curr Genet 1998, 33:368-377.

[23] Bryant, M. K. Functional analysis of genes encoding hydrolytic enzymes in the interaction of Epichloë festucae with perennial ryegrass. In PhD Palmerston North, NewZealand: Massey University; 2005.

[24] Tanaka, A., Tapper, B. A., Popay, A., Parker, E. J, Scott, B. A symbiosis expressed non-ribosomal peptide synthetase from a mutualistic fungal endophyte of perennial ryegrass confers protection to the symbiotum from insect herbivory. Mol Microbiol 2005, 57:1036-1050.

[25] Sambrook, J., Russell, D. W. Molecular Cloning: a laboratory manual. New York: Cold Spring Harbor Laboratory Press; 2001. BMC Evolutionary Biology 2009, 9:168

[26] Sanger, F., Nicklen, S., Coulson, A. R. DNA sequencing with chainterminating inhibitors. Proc Natl Acad Sci (USA) 1977, 74: 5463-5467.

[27] Pearson, W. R, Lipman, D. J. Improved tools for biological sequence comparison. Proc Natl Acad Sci U S A. 1988, 85(8): 2444-2448.

[28] Altschul, S.F, Gish, W., Miller, W, Myers, E. W, Lipman, D. J. Basic local alignment search tool. J Mol Biol 1990, 215:403-410.

[29] Altschul, S. F, Madden, T. L, Schaffer, A. A, Zhang, J, Zhang, Z, Miller, W, Lipman, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997,
25:3389-3402.

[30] Bendtsen, J. D, Nielsen, H, von Heijne, G, Brunak, S. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 2004, 340:783-795.
[31] Thompson, J. D, Higgins, D. G, Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994, 22:4673-4680.

[32] Tamura, K, Dudley, J, Nei, M, Kumar, S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 2007, 24:1596-1599.

[33] Guindon, S, Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology 2003, 52:696-704.

[34] Saitou, N, Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1987, 4:406-425.

[35] Felsenstein, J, Confidence limits on phylogenies: an approach using the bootstrap. Evolution 1985, 39:783-791.

[36] Zuckerkandl, E, Pauling, L. Evolutionary divergence and convergence in protease. In Evolving Genes and Proteins Edited by: Bryson V, Vogel HJ. New York: Academic Press; 1965:97-166.

[37] Saeki, K, Okuda, M, Hatada, Y, Kobayashi, T, Ito S, Takami, H, Horikoshi, K. Novel oxidatively stable subtilisin-like serine proteases from alkaliphilic Bacillus spp. enzymatic properties, sequences, and evolutionary relationships. Biochem Biophys Res Commun 2000, 279: 313-319.