Volume : 4 | Issue : 6 | June 2015 • ISSN No 2277 - 8179

Research Paper

Comparative Genome Analysis Using the Technique of Subtractive Hybridization

Biotechnology

KEYWORDS : Subtractive Hybridization, tester, driver, comparative genomics, gene expression.

Dr Suma Sarojini

Assistant Professor, Department of Biotechnology, Hosur Road, Christ University, Bangalore.

ABSTRACT

The era of genomics has provided us with a multitude of whole genome sequence data. The major task which lies ahead is to compare these genomes and find meaningful information regarding strain and species differences. This in turn can play a substantial role in the field of epidemiology and evolution. Differential gene expression analysis is very much required to study host specificity, virulence, habitat adaptation, stress response, secondary metabolite production etc. This review article is an attempt to elaborate the significance of Suppressive Subtractive Hybridization (SSH) which offers a handy solution to give preliminary information on the genes involved in these aspects. Although quite a number of sophisticated and costly techniques have been developed in this field, SSH still remains a simple and cheap method of comparing genomes especially when the whole genome sequence data is not available.

Introduction

Ever since the successful completion of the first whole genome sequence of the bacterium, Haemophilus influenzae (Fleishmann et al., 1995), a lot of genome sequencing works are being done all over the world. This has generated a huge amount of sequence data which has to be well studied and understood in order to give meaningful information and assign biological functions to the chunks of DNA. Comparative genomics is the field of biology in which the genome sequences of individuals within a genus or strains within a species are compared. There are many techniques to compare genomes. One of the simpler and cheaper ones is subtractive hybridization, which can be successfully utilized to get first hand information about the little differences between related species or strains. This essentially finds out unique DNA or RNA fragments present in the test species (tester) which are absent in the reference species (driver).

Need for comparative genomics

The need for comparative genomics arose from the availability of enormous data from various whole genome sequencing projects from mid 1990s. By comparing the completed genome sequences of one species with those of related organisms, researchers can identify regions of similarity and difference. This information will help researchers in studying various biological mechanisms. The potential ones include identifying conserved genes, understanding of differential behavior of related species or strains, evolutionary advantage of one species over the other, understanding gene insertions and deletions and figuring out strategies to identify drug targets and possible vaccine candidates. So ultimately the data obtained can be made use of to improve human health by combating diseases.

Genetic variation in bacteria is due to differences in gene content and nucleotide variations in or between structural genes. Small differences in DNA sequences, such as base changes in promoter or protein-coding sequences can change the timing or activity of a gene product and result in very different phenotypes between individuals. Differences in gene content are caused primarily by gene deletion or by acquisition of mobile elements such as plasmids and bacteriophages (Ochman et al., 2000). Understanding DNA sequence variations allows us to know the genetic basis of evolution, the genetic control of development, as well as the physiological abnormalities that result in disease. Detecting these sorts of variations has been very difficult due to the complexity of the genomes being analyzed and the small size of the differences. Those methods, which help to evaluate genome variability without full genome sequencing, are required to explain the evolutionary pattern and also to develop new effective therapeutic and diagnostic tools.

Subtractive hybridization technique

In the past, subtractive hybridization (SH) has been used with some success to identify large differences between two genomes, such as insertions or deletions. SH involves hybridizing DNA containing the sequence of interest with large amounts of the DNA lacking these sequences. The species carrying unique DNA fragments is termed as tester and the reference species is called driver in SH studies. Hybridization is usually followed by the physical separation of the undesired sequences from the target sequences. However, the larger and more complex the genomes or the smaller the difference between them, the more difficult this separation becomes. Successful subtractive hybridization usually results in only a 10- to 100-fold enrichment of the target sequences and therefore must be followed by the time-consuming process of sorting through large numbers of DNA sequences to find the one of interest (Myers et. al., 1993). The technique of representational difference analysis (RDA) couples subtractive hybridization with PCR to amplify exponentially only the target DNA sequences after repeated rounds of subtractive hybridization (Lisitsyn et al., 1993). The subtracted DNA sequences can then be cloned into vectors and sequenced. This can later be used to find upstream and downstream regions of the subtracted DNA, thereby mapping it to the genome.

While differences in the pathogenic potential may result from differential expression of proteins, the situation is more generally found to involve pathogen specific DNA sequence. Attributes of bacterial virulence are often grouped in islands and are frequently passed horizontally between two more closely related species. Representational difference analysis or RDA provides a quick means of cloning DNA corresponding to such species-specific sequence by direct physical subtraction of the chromosomal DNA of the pathogen. Unlike microarrays, the strength of this technique is its ability to identify regions that are present in some members of a species but absent from the genome of the sequenced strain. The identification of genetic variation amongst pathogenic bacterial strains is a good strategy for the discovery of strain-specific virulence-associated factors. Suppressive subtractive hybridization (SSH) is an easy and powerful tool for the identification of DNA fragments that are present in one strain (tester) but absent from another (driver). This technique is especially useful for the identification of genetic differences between two closely related bacteria whose genomes have not been sequenced.

Applications of Suppressive Subtractive Hybridization

1. Identification of strain or clone specific markers

One of the applications of SSH in the field of microbial genomics is the identification of strain or clone specific markers. Subtractive hybridization between the genomic DNA of M. avium strain 104 and M. intracellulare revealed 21 genes present in M. avium

that were absent in M. intracellulare, some of which coded for potential virulence factors that are unique to M. avium that may allow it to invade the intestinal mucosa more efficiently than M. intracellulare (McGarvey et al., 2001). Among the three phylogenetic lineages of Group B Streptococcus, strain III-3 is the one with most neonatal invasive disease. Subtraction has identified nine sequences specific to this strain, which may account for the difference in virulence (Bohnsack et al., 2001). Genomic subtraction between a pandemic Vibrio parahaemolyticus and a nonpandemic strain led to the identification of a 16-kbp insertion sequence present in almost all the pandemic strains but not in the nonpandemic strains tested (Okura et al.,2005). In the field of metagenomics, SSH has been used as a comparative method to examine the microbial diversity (i.e., species composition) and functional differences (i.e., gene composition) in the genomic content of two different rumen environmental communities (Galbraith et al., 2008).

Subtractive hybridization has also helped to identify mobile genetic elements that vary between strains. IS1598, an insertion sequence specific to virulent strains of Porphyromonas gingivalis, a potent pathogen in periodontal disease was identified by subtraction. (Sawada et al., 1999). Three clones with sequences similar to transposable elements specific to virulent Burkholderia psudomallei were generated when it was subtracted with avirulent B. thailandensis (Brown et al., 2000).

2. Identification of pathogenicity islands

The most interesting application of subtraction is the identification of genomic islands present in virulent strains, but absent from related avirulent strains. Subtractive hybridization followed by cloning of the end products was used to identify unique chromosomal regions associated with virulence of an E. coli strain (Brown et al., 1996). Subtracted clones were shown to pinpoint chromosomal regions shared by Neisseria meningitidis and Neisseria gonorrheae which are involved in the common aspects of host interaction of both pathogens and which are absent in the related commensal species N. lactamica. (Perrin et al., 1999). The larvae of Paenibacillus causes American Foulbrood of honeybees which is a fatal brood disease. SSH has helped in the identification of putative toxins and enzymes involved in antibiotic biosynthesis differing between four different genotypes of the species - ERIC I to IV, which differ in phenotype and also in virulence (Fünfhaus et al., 2009).

3. Comparison of gene expression

Subtractive hybridization is one of the simplest methods of comparing gene expression between species or strains of an organism. A large number of studies have been successfully done in the case of both prokaryotes and eukaryotes to this effect. Mycobacterium avium grown in human derived macrophage culture was subtracted with those grown in broth. This approach identified those genes expressed specifically during growth within macropahages (Plum et al., 1994). The two component regulatory system devR-devS found in M. tuberculosis and M. bovis and not seen in avirulent strains was identified by subtraction (Dasgupta et al., 2000). A study of genes involved in the pathogenicity mechanisms of Fusarium graminearum identified four genes (Abc2, Lyp1, Rrr1 and Zbc1) potentially involved in the pathogenesis and development of this fungus in wheat (Goswami et al., 2006).

SSH has significantly contributed in understanding plant gene expression too. A full-length cDNA SSH (FLSSH) method based on SSH and SMART cDNA synthesis was used to obtain differentially-expressed genes of rice treated by Probenazole (Yu et al., 2010). In a study of the drought response of wheat grass (Leymus secalinus), 16 sequences were found to be differentially expressed in response to drought (Yang et al., 2012). cDNA subtraction identified 73 genes to be up- or downregulated by ABA in pepper seedlings subjected to chilling stress (Wei et al., 2013).

4. Transcriptome subtractive hybridization (TSH)

TSH is a procedure which neither use restriction endonuclease for adapter ligation, nor use oligo-d (T) for the reverse transcription reaction. Instead, tester RNA is made to directly hybridize with driver single stranded cDNA. This hybrid of RNA/cDNA is digested by HaeIII, and the residual targeting tester RNA is enriched by reverse transcription PCR amplification. TSH has the advantage of covering different types of RNA as much as possible and reduce some of the experimental procedures (Ma et al., 2011).

Conclusion

This review explains the benefits of Suppressive Subtractive Hybridization in comparative genomic analysis. SSH being a simpler and cheaper tool, doesn't require much sophisticated instruments and is one of the efficient means of comparing two related genomes. Even without the whole genome sequence data, one can perform SSH. The advantage of SSH over microarrays lies in the fact that the latter cannot pull out sequences which is absent in the driver or reference strain. Even rare transcripts are detected by amplification in SSH. One of the disadvantages of this technique is the requirement of high quality RNA. This can be overcome by using high quality RNA isolation kits which can give high purity RNA for such work. A lot of modifications have been introduced in SSH thereby making it more reliable in the successful identification of unique genomic or cDNA regions. This is depicted in the results of a multitude of SSH experiments conducted both in prokaryotes and eukaryotes.

REFERENCE

1. Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A. & Kirkness, E. F. (1995). Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science 269: 496-512. | 2. Ochman, H., & Lawrence., Groisman, E, A., (2000). Lateral gene transfer and the nature of bacterial innovation. Nature. 405 (6784): 299-304. | 3. Myers, R, M. (1993). The pluses of subtraction. Science 259:942-943. | 4. Lisistsyn, N. A., Lisistsyn, N, M and M, Wigler. (1993). Cloning the difference between two complex genomes. Science 259: 946–51. | 5. Bohnsack, J, F., Takahashi, S., Detrick, S, R., Pelinka, L,R., Hammitt, L,L., Aly, A., Whiting, A and Adderson, E, E. (2001). Phylogenetic classification of serotype III group B streptococci on the basis of hylB gene analysis and DNA sequences specific to restriction digest pattern type III-3. J. Infect. Dis. 183:1694-1697. | 6. McGarvey, J. A and Bermudez. L, E. (2001). Phenotypic and genomic analyses of the Mycobacterium avium complex reveal differences in gastrointestinal invasion and genomic composition. Infect. Immun 69:7242-9. | 7. Okura, M., Osawa, R., Arakawa, E., Terajima, J & Watanabe, H. (2005). Identification of Vibrio parahaemolyticus Pandemic Group-Specific DNA Sequence by Genomic Subtraction. J. Clin Micro. 43(7): 3533-36. | 8. Galbraith, E, A., Antonopoulos, D, A & White, B. (2008). Application of suppressive subtractive hybridization to uncover the metagenomic diversity of environmental samples. Methods Mol. Biol.410:295-333. | 9. Sawada, K.,, Kokeguchi, S.,, Hongyo, H., Sawada, S., Miyamoto, M., Maeda, H., Nishimura, F., Takashiba, S & Murayama, Y. (1999). Identification by subtractive hybridization of a novel insertion sequence specific for virulent strains of Porphyromonas gingivalis. Infect. Immun .67(11):5621-25. | 10. Brown, N.F., Lew A, E & Beacham, I, R. (2000). Identification of new transposable genetic elements in Burkholderia pseudomallei using subtractive hybridization. FEMS Microbiol. Lett. 183:73-79. | 11. Brown, P. K and Curtiss, R. (1996). Unique chromosomal regions associated with virulence of an avian pathogenic E. coli strain. Proc. Natl.Acad. Sci. U.S.A. 93: 11149-11154. | 12. Perrin, A., Nassif, X & Tinsley, C. (1999). Identification of regions of the chromosome of Neisseria meningitidis and Neisseria gonorrhoeae which are specific to the pathogenic Neisseria species. Infect Immun. 67(11): 6119-29. | 13. Fünfhaus, A., Ashiralieva, A., Borris, R & Genersch, E. (2009). Use of suppression subtractive hybridization to identify genetic differences between differentially virulent genotypes of Paenibacillus larvae, the etiological agent of American Foulbrood of honeybees. Environ. Microbiol. 1(4):240-50. | 14. Plum, G. and J. E. Clark-Curtiss. (1994). Induction of Mycobacterium avium gene expression following phagocytosis by human macrophages. Infect. Immun. 62:476-483. | 15. Dasgupta, N., V. Kapur, K. K. Singh, T. K. Das, S. Sachdeva, K. Jyothisri, and J. S. Tyagi. (2000). Characterization of a two-component system, devR-devS, of Mycobacterium tuberculosis. Tuber. Lung Dis. 80:141-59. | 16. Goswami, R, S., Xu, J, R., Trail, F., Hilburn, K & Kistler, H, C. (2006). Genomic Analysis of Host-Pathogen Interaction between Fusarium graminearum and Wheat during Early Stages of Disease Development. Microbiology, 152 (6): 1877-1890. | 17. Yu, Z & YaoGuang, L. (2010). A full-length cDNA subtractive hybridization (FLSSH) method and its application in the isolation of inducible expression genes in rice. Journal of South China Agricultural University 31 (1): 34-37. | 18. Yang, M. Jiang, Y. Cai, Y. Ye and W. Lihuan, W. (2012). Construction and Analysis of Suppression Subtractive Hybridization Library for Leymus secalinus Exposed to Drought Stress. African Journal of Biotechnology.11 (73): 13775-13781. | 19. Guo, W, L., Chen, R, G., Gong Z, H., Yin, Y, X., & Li D, W. (2013). Suppression Subtractive Hybridization Analysis of Genes Regulated by Application of Exogenous Abscisic Acid in Pepper Plant (Capsicum annuum L.) Leaves under Chilling Stress. PloS One 8(6): e66667. | 20. Ma, Y., Hong Z., Wu, Q & Yao X. (2011). Transcriptome Subtractive Hybridization and its Reliability Validated by an E. coli Model. Bioscience Methods, 2 (6). (doi:10.5376/bm.2011.02.0006). |