ready be applied for sequence analysis and checking the accuracy of conventional sequencing methods. A number of other applications in the Human Genome Program are within the reach of SHOM, such as mutation screening, gene polymorphism studies, detection of microorganisms, gene expression studies, etc. Application of SHOM for de novo DNA sequencing requires manufacturing of more complicated microchips and improvement of some other, already available methods.

DOE Contract No. W-31-109-Eng-38.

References

1. Yershov G., Barsky V., Belgovsky A., Kirillov Eu., Kreindlin E., Ivanov I., Parinov S., Guschin D., Drobishev A., Dubiley S., Mirzabekov A. DNA analysis and diagnostics on oligonucleotide microchips. // Proc. Natl. Acad. Sci. 1996. Vol. 93. 4913-4918.
2. Parinov S., Barsky V., Yershov G., Kirillov Eu., Timofeev E., Belgovskiy A., Mirzabekov A. DNA sequencing by hybridization to microchip octa-and decanucleotides extended by stacked pentanucleotides. // Nucl. Acids Res. 1996. Vol. 24. N 15. P. 2998-3004.
3. Timofeev E., Kochetkova S., A., Mirzabekov A. Radioselective immobilization of short oligonucleotides to acrylic copolymer gels // Nucl. Acids Res. 1996. Vol. 24. N 16. P. 3142-3148.
4. Prudnikov D., Mirzabekov A. Chemical methods of DNA and RNA fluorescent labelling. // Nucl. Acids Res. 1996., in press.
5. Livshits M., Mirzabekov A. Theoretical analysis of the kinetics of DNA hybridization with gel-immobilized oligonucleotides. // Biophys. J. 1996. Vol. 71, in print//.
6. Guschin D., Mobarry B., Proudnikov D., Stahl D., Rittmann B., Mirzabekov A. Oligonucleotide microchips as genosensors for determinative and environmental studies in microbiology //Applied and Environmental Microbiology, in print//.
7. Guschin D., Yershov G., Zaslavsky A., Gemmell A., Shick V., Lysov Yu., Mirzabekov A. A simple method of oligonucleotide microchip manufacturing and properties of the microchips // submitted for publication.
8. Drobyshev A., Mologina N., Shik V., Pobedimskaya D., Yershov G., Mirzabekov A. Sequence analysis by hybridization with oligonucleotide microchip: identification of beta-thalassemia mutations // Gene (in print).
9. Dubiley S., Kirillov Eu., Lysov Yu., Mirzabekov A. DNA fractionation, sequence analysis and ligation of immobilized oligomers on oligonucleotide chips // submitted for publication.
10. Timofeev E., Smirnov I.P., Haff L.A., Tishchenko E.I., Mirzabekov A.D., Florentiev V.L.. Methidium Intercalator Inserted into Synthetic Oligonucleotides // Tetrahedron Letters 1996, v.37, N47, p.8467.

Relevant Publication

Methods of DNA sequencing by hybridization based on optimizing concentration of matrix-bound oligonucleotide and device for carrying out same by Khrapko K., Khorlin A., Ivanov I., Ershov G., Lysov Yu., Florentiev V., Mirzabekov A. US Patent 5,552,270, Sep. 3, 1996. PCT/RU92/00052, filed Mar 18, 1992.

High-Throughput DNA Sequencing: SAmple SEquencing (SASE) Analysis as a Framework for Identifying Genes and Complete Large-Scale Genomic Sequencing

Robert K. Moyzis and Jeffrey K. Griffith[1]
Center for Human Genome Studies; Los Alamos National Laboratory; Los Alamos, NM 87545
505/667-3912, Fax: -2891, *moyzis@telomere.lanl.gov*
[1]University of New Mexico; Albuquerque, NM 87131

The human chromosome 5 and 16 physical maps (Doggett et al., Nature 377:Suppl:335-365, 1995; Grady et al., Genomics 32:91-96, 1996) provide the ideal framework for initiating large-scale DNA sequencing. These physical mapping studies have shown clearly that gene density in humans will vary greatly. For example, band 16q21, consisting of 8 Mb of DNA, has no genes or trapped exons assigned to it, as yet. In contrast, band 16p13.3 has an extremely high density of coding regions in the DNA examined to date (i.e., multiple genes/cosmid). Given this wide variation in gene density and current sequencing costs, we propose that newly targeted genomic regions should be analyzed first by a "Lewis and Clark" exploratory approach, before committing to full length DNA sequencing. We are using a SAmple SEquencing (SASE) approach to rapidly generate aligned sequences along the chromosome 5 and 16 physical maps. SASE analysis is a method for rapidly "scanning" large genomic regions with minimal cost, identifying, and localizing most genes. Briefly, individual cosmids are partially digested with Sau3A and 3 kb fragments are recloned into double-strand sequencing vectors. By sequencing both ends of a 1X sampling of these recloned fragments along with end sequences of the cosmid, 70% sequence coverage is achieved with 98% clone coverage. The majority of this clone coverage is ordered by the relationship between the subclone end sequences. These ordered sequences are ideal substrates for directed sequencing strategies (for example, primer walking or transposon sequencing). SASE analysis has been initiated on the 40 Mb short arm of chromosome 16 and the 45 Mb short arm of chromosome 5. We propose to make SASE sequences, along with feature annotation, publicly available through GSDB. Such data are sufficient to allow PCR amplification of the sequenced region from GSDB submissions alone, eliminating the need for extensive clone archiving and distributing, will allow for the effective "democratization" of the genome, allowing numerous laboratories to share and contribute to the growing genome databases.

DOE Grant No. DE-FG03-96ER62298.