RESEARCH ARTICLE

Automated Large-Scale Genome Scanning and Linkage Analysis Strategy of Genetic Disease Related Gene Location

Lawrence Noyes, Chiara Palacios, Reginald Finn

Department of Public Nursing and Caring, Sciences and Technology University, Philadelphia

Abstract: With the development of the Human Genome Project, many methods have been applied in this field. Since the 1990s, large-scale scanning and mapping of base circles have been established and developed by combining linkage analysis with automation in genomes. At present, it has become the most widely used and effective means in this field.

Keywords: Linkage analysis; genetic disease; gene mapping PCR

Since the human genome project was launched in the early 1990s, great progress has been made in this field. The four major objectives of the Human Genome Project are to complete the genetic map and physical map of the human genome. The genetic map has been completed ahead of schedule. At present, more than 6,000 ideal markers have been isolated, covering an average area of 0.7 cM, which has met the requirements of the genetic map. The physical map is drawing to a close. The sequence map will be sequenced by 2005. At present, the construction of expression maps is also developing rapidly. More than one million fragments of cDNA have been cloned. In these studies, the construction of the expression map is particularly noticeable. Many genes related to hereditary diseases have been mapped, such as obesity (0), rational fibrosis (CF), Huntington's chorea (HD), Alzheimer's disease (AD), hereditary ataxia (FRDA), tel-angiectasia syndrome (AT), and so on. Other genes related to hereditary diseases have also been mapped. More progress, such as diabetes. A variety of theories and techniques in biology, computer science, mathematics and other disciplines have been applied to this field, especially the rapid development of genetic and physical maps. A method based on linkage analysis to scan, typing and locate related genes in genome automatically and on a large scale has become this method. A powerful tool in the field of research[2]. This article will briefly review the principle and technical process of this method.

1. Basic principles

1.1 Linkage analysis, sequence diagram and expression map.

Location based cloning strategy based on linkage analysis is one of the main strategies in gene mapping. This is especially true when genetic products are known. The basic principle of linkage analysis is that in a family, two loci on the same chromosome (pathogenic genes and genetic markers) exchange and recombine during meiosis. The higher the recombination rate, the less chance the two loci will pass on to their offspring together. By genotyping the marker in the genetic map with appropriate coverage density in the family, a marker closely linked to the pathogenic gene can be found, thus determining the size of the gene on the chromosome. On this basis, the disease-causing gene can be identified in a smaller region by further linkage analysis using a higher-coverage marker in the chromosomal region, and then the gene can be identified or cloned using the dye block fragment of the region provided by the physical map. The com-

Copyright © 2018 Lawrence Noyes *et al*.
doi: 10.18063/gds.v2i1.
This is an open-access article distributed under the terms of the Creative Commons Attribution Unported License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

pletion of genetic maps and the development of physical maps provide the basis for this analysis. The invention of molecular biology techniques such as PCR and the use of computers and automatic sequencers have now automated this technique.

1.2 DNA polymorphic markers have many markers for gene mapping in genetic diseases.

Species, such as traditional ABO blood group system, serum protein diversity, HLA antigen and so on. Because these systems have too few loci on DNA to cover the entire genome well, their applications in gene mapping are limited and are now largely unused.

1.2.1 RFLPs (restriction fragment length polymorphisms)

A restricted fragment length polymorphism method was established for the mid and late 70s. The number of loci determined by this marker in the whole genome is more than 105. Once established, the system is widely applied to genome research. But its detection needs Southern Blot and other processes, and the steps are cumbersome. Although the invention of PCR can improve the detection speed, but eventually the discovery of more superior polymorphic markers makes RFLP method gradually eliminated.

1.2.2 VNTRs (variable number of tandem repeats)

The method is called multiplicity quantity random repeat sequence marker. In 1985, Alec Jeffreys of Leicester University and his colleagues discovered short, simple repeating units in the myosin gene, which he called the "minisatellite core"[3]. Subsequent studies have gradually shown that a large number of polymorphic markers of this length are distributed in the human genome. In 1989, another class of markers, microsatellite marker system, was used in positional cloning. These markers were shown to be 2, 3 or 4 tandem repeats of nucleoxic acid. Because of their wide distribution in the genome, large number of alleles and Mendelian inheritance, they can provide enough genetic information for linkage analysis and are relatively easy to detect by PCR and electrophoresis. This method is currently the most widely used marker in gene mapping research. System.

2. Automatic linkage analysis

2.1 Gene mapping is generally used in the collection and processing of families. As an ideal family for linkage analysis, it should be strictly diagnosed, positive and require a relatively closed population with small migration; the number of days and the size of the family should meet certain requirements; and the family should have a clear legacy. Pass the relevant data, such as heritability, Xi, and so on. After collecting blood samples from families, DNA is extracted in a rigorous manner. Many laboratories now use specialized kits for DNA extraction and purification.

2.2 The selection of microsatellite markers for preparation of microsatellite markers should generally reach the average coverage of 5-lOcM on chromosomes. These microsatellite markers were then synthesized using PCR primers, one of which was labeled with fluorescent substances (such as FITC) during the synthesis process.

The PCR reaction is carried out in 96 well plates. The preparation of PCR reaction solution can be completed by sampling and adding samples from other machine tools.After the PCR reaction is completed, the product can be used for electrophoresis detection by adding "termination solution" such as methyl-cool glue, and then heating to 95 degree C to denaturate the product.

One of the current trends is that primers of microsatellite markers are labeled with several different fluorescent substances, and multiple PCR (multiplex PCR) is performed in the same reaction tube to make the PCR process more rapid and convenient.

2.3 Electrophoresis detection

The detection of PCR products is carried out on automatic excitation fluorescence sequencer (ALF) (Pharmacia, PE and other companies). In the gel electrophoresis, the special laser probe on the sequencer can excite the molecular weight of the PCR products that the well scans for each swimming lane. The test can detect multiple PCR products on each swimming lane, and the glue can be reused.

2.4 Automatic genotyping

The computer system connected with the sequencer can intuitively express the yield and molecular weight of each product by curve peak spectrum, and the data detected by electrophoresis can be processed by corresponding software system. First, the software accurately measures the molecular weight internal standard in each lane, and then draws the regression curve of molecular weight based on the internal standard. Second, the software further determines the authenticity of each peak. Due to the uncertainties in the PCR reaction, some "noise flying" peaks, commonly known as "stut ter" and "shadow" peaks, appear on the electrophoresis spectrum. The reason is not very clear. These peaks can be removed by computer after comprehensive calculation and analysis; in addition, the repetitive markers of tetranucleotide have a low general harmonic noise[4] and are now used more; and in-depth detection of each product to determine the molecular weight of each product according to the regression state line, and thus Fourth, the genetic model of the product was checked and checked. Fifth, the genotype of each product was determined and stored in the software for further analysis. Sixth, the data after the above processing were input into LIKAGE and other software for LOD calculation.

2.5 Genome scan comprehensive analysis

The genetic LOD value of each microsatellite marker can be calculated by combining the data processed above with the data obtained from the family analysis. If the LOD value of a marker is high (generally considered to be more than 3 affirmative linkage), the marker can be considered to be linked to the pathogenic gene. By this method, the genetic LOD value of each microsatellite marker can be calculated. The pathogenic gene can be located in a certain segment of a chromosome. Further linkage analysis by selecting more dense loci and enlarging the number of families in this region could identify the gene in a narrower region, which could be identified by sequencing.

This approach has played an important role in the study of genetic diseases, especially monogenic diseases, such as Wolfram syndrom and Freiderich Ataxia. It also has great potential in the study of complex genetic diseases such as schizophrenia, manic depression, essential hypertension, diabetes and so on. For example, the team's Herring researchers in the United Kingdom used this method to locate the gene for manic depression on chromosome 4 on the short arm[5]. Science and other journals have reported on this[6] 0 At present, our group is working on the use of this system for gene mapping in some psychiatric and neurogenetic diseases.

3. Prospects

With the development of the Human Genome Project (HGP), the "post-genome project" has been proposed to study gene mapping and function. The orientation, cloning and development of various genes, especially the genes of genetic diseases, will become the mainstream of genetic research in the future. China has a vast and densely populated area, a large number of ethnic groups, less migration, relatively isolated skin, a fifth of the world's population, but also has the world's most abundant and ideal genetic family resources. It is of great significance to make full use of this method to promote the research of genetic diseases and improve the level of science and technology in China. At present, this method has been developed and combined with other methods such as association analysis and linkage disequilibrium analysis, which greatly improves the ability of gene mapping in polygenic and complex genetic diseases. It is believed that this method will be continuously promoted and developed in the genetic research of genetic diseases, especially in China.

4. The relationship between heredity and human health

Human health is the result of a combination of many factors. Among human health factors, genetic factors account for 10%-15%. According to statistics, 15% of the world's population is genetically ill. Some high mortality, high risk and refractory diseases are genetically related. For example, cancer, cardiovascular disease, hypertension, diabetes and so on. More and more new techniques have been applied to the analysis of chromosomal abnormalities, resulting in thousands of inherited diseases. It is concluded that heredity is an important factor affecting human health.

Metabolism is the basic characteristic of organisms. It maintains and sustains life by taking nutrients from the surrounding environment and transforming them into energy that can be used by oneself. Specific metabolic patterns de-

pend on the particular genetic structure of organisms. Therefore, human health, in a sense, is the result of the balance between metabolism determined by the original genetic structure of the human body and the environment in which the human body is located. Both changes in genetic structure and changes in the surrounding environment will result. Break the balance between the normal metabolism and the surrounding environment, resulting in the occurrence of diseases, affecting human health.

4.1 The harm of heredity to human health

At present, more than 4,000 genetic diseases have been discovered in the world. This number is increasing at the rate of adding 100 new diseases every year. Most of these genetic diseases are common and frequently-occurring. In our country, 15 million of the 300 million children suffer from genetic diseases. These congenital diseases and inherited diseases of mental retardation not only bring great pain to the body and mind of patients, but also bring burden to their families and society.

Hereditary diseases are affected by individual genetic genes, which are transmitted among members of the blood relationship. Some affected individuals have pathogenic genes as early as the fertilized egg stage. This gene will not only accompany the life of the patient, but also be passed on to their own children and offspring through inheritance. Because some diseases are dominant, easy to discover, easy to treat and control, can take certain measures to prevent the disease of childbearing after marriage, and some pathogenic genes are recessive, not easy to find, for prevention and treatment have brought great difficulty, also for human health and society. Development has brought greater harm.

References

1. Song Li, Zhang Xiuqing, Zheng Naiqi, *et al*. Foreign medical genetics, 1997, 20 (2): 102--105
2. He Lin Mansfield D C, Brown A F, *et al*. et, Lin, 1995
3. Jeffereys A, fly called Tilson V, Thein S. Nature, 1985; 314: G7 73
4. He Lin, Morris S, Lennon A, *et al*. Psychia Genet Psychia 123-129
5. Blackwood D H R He Lin Morris S, *et al*. et, 1996; 12 (4): 427-430
6. Morell V. Science, 1996; 272: 31-32
7. Mou Xin.Preliminary study on the correlation between the essence of TCM syndrome and epigenetics [J].Zhejiang Journal of Integrated Chinese and Western Medicine, 2013 (6): 443-446.
8. Ma Teng, Xu Jing, Yang Yajun. Genetic polymorphism of 15 X-STR loci in Shandong Han population [J].Forensic Journal, 2013 (3): 202-205, 208.
9. Meng Zong Jun. The guiding strategy of the research on human genetic traits [J]. Guangxi education.2013 (10):21-22.
10. Yan Rui. Cardiovascular disease is the leading cause of premature death in schizophrenics.
11. Chinese Journal of Evidence-based Cardiovascular Medicine, 2013 (1): 61.