SECOND SOUTH - NORTH HUMAN GENOME CONFERENCE

November 6-10, 1994
Beijing, China

United Nations Educational, Scientific and Cultural Organization (UNESCO)
Chinese National Commission for UNESCO
Peking University
Preface

As the most important international collaborative project in life sciences, the Human Genome Program is attracting the interest and attention of many scientists and their governments. In order to promote progress in this field, especially in developing countries, the First South-North Human Genome Conference was held in Brazil in 1993.

My colleagues and I at Peking University were greatly honored and delighted to be requested by UNESCO and the Chinese National Commission for UNESCO to help organize the Second South-North Human Genome Conference which took place in Beijing, China during the period of November 6-10, 1994.

About 30 leading scientists from sixteen countries, including the Nobel laureate Dr. C. Gajdusek, attended the Conference and together with their Chinese counterparts presented more than 40 papers distributed in 6 sessions. The Conference provided an excellent opportunity for the participants to share their experiences and expertise, thereby serving the same purpose of promoting international exchange and progress in this very exciting and fast-moving field.

We were honored at the opening and closing ceremonies by the presence of Director-General of UNESCO Dr. F. Mayor, Chairman of the International organizing Committee Dr. S. Grisolia, State Councilor Dr. Jian Song, Chairwoman of the Chinese National Commission for UNESCO Dr. Yu Wei, Director of the Chinese National Science Foundation Dr. Cun-hao Zhang, Minister of Public Health Dr. Ming-zhang Chen, and many other leading Chinese authorities.

I would like to take this opportunity to express my sincere gratitude to UNESCO, to the Chinese National Commission for UNESCO, to our two organizing committees, the participants and the secretariat for their support and enthusiasm which made this Conference successful. Special thanks go to Mrs. Matsui, Program Specialist in Biology, Basic Science Division of UNESCO, for her contribution as Chairwoman of the Liaison group.

Zhang-liang Chen, Ph. D
Professor, Vice President, Peking University
Co-Chairman, The Second South-North Human Genome Conference
Contents

Organizers and Sponsors ......................................................... 1

Co. Sponsors ............................................................................. 2

Local Organizing Committee ..................................................... 3

International Advisory Committee ............................................ 4

International Organizing Committee ......................................... 5

Liaison Group ........................................................................... 5

Timetable .................................................................................. 6

Map of Convention Center ........................................................ 7

Program ..................................................................................... 8

Summary of Presentations ......................................................... 14

An Address to the Second South-North Human Genome Conference
Song, Jian (Beijing, China) ......................................................... 15

Address by Mrs. Wei Yu
Wei, Yu (Beijing, China) ........................................................... 17

Address by Zhang, Cun-hao
Zhang, Cun-hao (Beijing, China) .............................................. 19

Address by Hu, Zhao-guang
Hu, Zhao-guang (Beijing, China) .............................................. 21
Address by Xu, Zhihong
Xu, Zhihong (Beijing, China) ................................................. 22

UNESCO Scientific Programme on the Human Genome
Grisolia, Santiago (Valencia, Spain and Kansas City, USA)..................... 24

The Chinese Human Genome Initiative
Qiang, Boqin. MD (Beijing, China)............................................. 27

The Distribution of Genes in the Human Genome
Bernardi, G. (Paris, France)...................................................... 33

Map-driven Gene Finding to Fulfill the Genome Project
Schiessinger, David (Saint-Louis, USA)........................................ 37

YAC-Based Mapping of Human Xq26-qter by Probe and STS content
D’Urso, Michael (Naples, Italy). ............................................... 42

Esophageal Cancer Disease Genes
Wu, Min (Beijing, China)......................................................... 47

Genes Causing Congenital Malformations and Heritable Cancer
Syndromes: The Example of the RET Protooncogene
Romeo, Giovanni (Geneva, Italy)............................................. 51

Molecular Genetics of Cystic Fibrosis and Other Diseases Associated with
Human Chromosome 7
Tsui, Lap-Chee (Toronto, Canada)........................................... 54

Positional Cloning of X-linked Disease Genes
Monaco, A. P. (Oxford, UK)...................................................... 58

Genes Related to Human Primary Hepatitis Cancer
Gu, Jian-ren (Shanghai, China) ...........................................62

Genes Involved in Acute Promyelocytic Leukemia (APL)
Chen, Zhu et al (Shanghai, China) .............................................66

Molecular Genetic Basis of Amyotrophic Lateral Sclerosis
Deng, Hanxiang (Hunan, China) ..............................................71

Human Genetic Diversity and Its Evolutionary Implication
Gojobori, T. (Mishima, Japan) .................................................74

Fourteen-Years Studies on the Genetic Diversity of Ethnic Groups in China
Du, Ruofu (Beijing, China) .....................................................78

Chinese Genome Project and Program Report of PKU Lab
Chen, Zhangliang (Beijing, China) ............................................82

Rice Molecular Mapping and Gene Location
Zhu, Lihuang et al (Beijing, China) ..........................................84

Twenty Five Ethnic Minorities in Yunnan Province: Their Genotypic and Phenotypic Characteristics
Chu, Jiayou (Beijing, China) ....................................................86

Analysis of Some Rice Gene Sequence and Analysis of Rice Chromosomes
Sun, Chong-rong et al (Shanghai, China) ..................................92

Sequencing, Identification and Mapping of Schistosoma mansoni Genes
Pena, S. et al (Brasil) ..............................................................96

Automated DNA Sequencer with a New Continuous Detection System for Simultaneous On-line Sequencing of 44 Clones
Ansorge, W. (Heidelberg, Germany) ........................................101

III
Overview of Genome Sequencing Strategies
Chen. Ellson Y. (Foster City, USA)................. 111

New Biology Needs New Approaches: DNA Sequencing by Hybridization and Protein Mapping on DNA by Crosslinking
Mirzabekov. A. (Moscow, Russian Federation)........ 115

Construction of Primers by Ligation of Short Degenerated Oligomers and Rapid Sequencing Without Conventional Cloning
Szybalski, W. (Madison, USA)......................... 117

Flourescent Differential Display: A High Throughput System for a Large-scale Scanning of the Differentially Expressed Genes
Sakaki, Y. et al (Tokyo, Japan).......................... 120

Development of Simple Molecular Biology Techniques for Genome Researching Latin America
Pena, S. et al. (Brasil)..................................... 125

New Concepts in Pathogenesis of Slow Viral Infections
Gajdusek, D. (Bethesda, USA)............................ 132

The Patient Recovers
Grey, E. Dimond (Missouri, USA)...................... 134

Medical Genetic Service in Human Genome Project
Fujiki, N. et al. (Fukui, Japan)......................... 136

Study on Physical Mapping of Human Chromosome 17q11-q12 by Construction of YAC Contigs
Han, Shunsheng, et al. (Shanghai and Human, China)...... 138

The Recent Progress of Animal Genome Mapping in China

IV
Li, Ning, et al. (Beijing, China) .................................................. 140

Study on Molecular Structure of Dystrophin Gene in Chinese
Long, Yu, et al. (Shanghai, China) .......................... 146

Analysis of Random Amplified polymorphic DNA for Four Minorities in Yunnan Province
Wang, wen, et al. (Kunming, China) .......................... 148

Identification of DNA Makers Linked to Blast Resistance Genes in Rice
Zheng, KL, Kochert, G, et al. (Hangzhou, China and Athens, USA) .......................... 150

The Search for Human Nodal Gene
Zhou, X., et al. (Beijing, China) .......................... 152

Closure Address on the South-North Human Genome Conference
Chen, Ming-zhang (Beijing, China) .......................... 153

Address at the Closing Meeting of the Second South-North Human Genome Conference
Zhang, Chong-li (Beijing, China) .......................... 155

List of Participants ................................................................. 157
ORGANIZERS
United Nations Educational, Scientific and Cultural Organization (UNESCO)
Chinese National Commission for UNESCO
Peking University

SPONSORS
State Science and Technology Commission of China
National Natural Science Foundation of China
Third World Academy of Sciences
State Education Commission of China
Co-sponsors

1. China Winning Biotech Group
2. Pharmacia Biotech (China) Ltd., Sweden
3. Wogen Technology Limited, England
4. New Brunswick Scientific Co. Inc. USA
5. Boehringer Mannheim China Ltd., Germany
6. Silicon Graphics, Inc. USA
7. Perkin-Elmer Applied Biosystems Division, USA
8. Line Analytics Ltd., Hong Kong
9. Cole-Parmer, USA
10. B. Braun Biotech International, Germany
11. Fortune on Instrument Co. Hong Kong
12. Bio-Rad, USA
13. Waters Chromatography Division, USA
14. Taike Instruments Ltd. Beijing, China
15. The Oriental Science & Technology Co., China
16. Promega Corporation, USA
17. Donglian Electronics Ltd., Harbin, China
18. Taicang Scientific Instruments, China
19. Academy of Military Medical Sciences
20. Millipore Corp. in China
Local Organizing Committee

Chairman
Chen. Z.-L. Dean and Professor. Peking University

Members
Wei, Y. Chairwoman, Chinese National Commission for UNESCO
Zhang, C.-H. Director, National Natural Science Foundation of China
Zhu, L.-L. Vice Chairwoman, State Science and Technology Commission of China
He, J.-S. Vice Minister, Ministry of Public Health of China
Xu, Z.-H. Vice President, Chinese Academy of Sciences (Academia Sinica)
Wu, S.-Q. President, Peking University
Liu, Y.-H. Director, China National Center for Biotechnology Development
Zhang, C.-L. Deputy Secretary-General, Chinese National Commission for UNESCO
Qiang, B.-Q. Vice President, Chinese Academy of Medical Sciences
Chen, Z. Professor, Shanghai Second Medical University
Chu, J.-Y Professor, Chinese Academy of Medical Sciences
Hong, G.-F. Professor, Chinese Academy of Sciences
Deng, H.-X. Professor, Hunan Medical University
Gu, J.-R. Professor, Shanghai Tumor Research Institute
Gu, H.-Y. Professor, College of Life Sciences, Peking University
International Advisory Committee

Chairman
Mayor. F.  Director-General of UNESCO

Members
Bernardi, G.  Professor, Laboratoire de Genetique Moleculaire, Institut Jacques Monod
Caskey, G.  Professor, Institute for Molecular Genetics
Dausset, J.  Professor
Fasella, P.  Professor, Commission of the European Communities
Grisolia, S.  Professor, Instituto de Investigaciones Citologicas
McKusick, V.  Professor, Johns Hopkins Hospital
Mirzabekov, A.  Professor, Academy of Sciences of Russian Federation
Ole-Moi Yoi, O  Professor, International Laboratory for Research on Animal Diseases
Pena, S.  Professor, Latin American Human Genome Program
Tan, C.-C.  Professor, Institute of Genetics, Fudan University
Wu, M.  Professor, Institute of Oncology, Chinese Academy of Medical Sciences
International Organizing Committee

Chairman
Grisolia, S. Professor, Instituto de Investigaciones Citologicas

Members
Bernardi, G. Professor, Laboratoire de Genetique Moleculaire, Institut Jacques Monod
Cai, R.-S. Director, Science Division, China National Commission for UNESCO
Chen, Z.-L. Professor, College of Life Sciences, Peking University

Liaison Group

Chairwoman
Mrs. Matsui, S. Program Specialist in Biology, Basic Science Division, UNESCO

Members
Bernardi, G. Professor, Laboratoire de Genetique Moleculaire, Institut Jacques Monod
Grisolia, S. Professor, Instituto de Investigaciones Citologicas
Gu, X.-C. Professor, College of Life Sciences, Peking University
Yuan, K.-W. First Secretary, Chinese Permanent Delegation to UNESCO
Zharov, V. Director, Basic Science Division, UNESCO
## Time Table

**Location: Room 3017, International Convention Center**

<table>
<thead>
<tr>
<th>Nov. 6</th>
<th>Morning</th>
<th>Afternoon</th>
<th>Evening</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sun</td>
<td>Registration</td>
<td>15:00 -- 18:00</td>
<td>Reception</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18:30</td>
</tr>
<tr>
<td>Nov. 7</td>
<td>Opening Session 10:00 -- 12:00</td>
<td>Session I Human Genome 13:30 -- 18:00</td>
<td>Peking Opera 19:00 -- 21:30</td>
</tr>
<tr>
<td>Mon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nov. 8</td>
<td>Session II Human Genetic Diseases 8:30 -- 12:00</td>
<td>Session III Human Diversity Session IV Other Genome Projects 13:30 -- 18:00</td>
<td>Banquet 19:00</td>
</tr>
<tr>
<td>Tue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nov. 9</td>
<td>Session V Technical Advances 8:30 -- 12:00</td>
<td>Session VI Genome Projects in Developing Countries 13:30 -- 15:30 Closing Ceremony 16:00</td>
<td></td>
</tr>
<tr>
<td>Wed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nov. 10</td>
<td>Sightseeing 8:30</td>
<td>Group 1: The Great Wall, Ming Tombs Group 2: Forbidden City, Temple of Heaven</td>
<td></td>
</tr>
<tr>
<td>Thu.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Optical Character Recognition (OCR) document. WARNING! Spelling errors might subsist. In order to access to the original document in image form, click on "Original" button on 1st page.
SECOND SOUTH - NORTH
HUMAN GENOME CONFERENCE

PROGRAM
November 6, 1994

15:00 - 18:00 Registration
18:00 Reception

November 7, 1994

10:00-12:30 Opening Session

Chairman: Prof. Zhang-liang Chen

1. Address by Dr. Jian Song,
   State Councilor and Chairman of State Science and Technology
   Commission of China
2. Address by Dr. Yu Wei, Chairwoman, Chinese National Commission
   for UNESCO
3. Welcome Speech by Mr. Qi-yan Li, Mayor of Beijing City
4. Address by Dr. Federico Mayor, Director-General of UNESCO
5. Addresses by Mr. Cun-hao Zhang (Director, NNSFC)
   Mr. Zhi-hong Xu (Vice President, CAS)

Break

Chairmen: Prof. D. Gajdusek and Prof. Min Wu

6. A gene map accelerates utility of the genome initiative
   T. Caskey, (Houston, USA)
7 UNESCO scientific program on the Human Genome
   S. Grisolia (Valencia, Spain and Kansas City, USA)
8. The Chinese Human Genome Initiative
   B. -Q. Qiang (Beijing, China)

13:00 - 18:00

Session 1 The Human Genome

Chairmen: Dr. Mckusick and Prof. B. -Q. Qiang
9. The distribution of genes in the human genome  
   G Bernardi (Paris, France)
10. Integration of the cytogenetic, genetic and physical maps of human  
   Chromosomes  
   D. Ward (New Haven, USA)
11. Map-driven gene finding to fulfill the genome project  
   D. Schlesinger (Saint-Louis, USA)

**Break**

12. YAC-based mapping of human Xq26-q28 by probe and STS content  
    M. D’Urso (Naples, Italy)
13. Compositional compartments of human genome and viral integration  
    A. Rynditch (Kiev, Ukraine)

**November 8, 1994**

8:30-12:00  
**Session 2 Human Genetic Diseases**

Chairmen: Dr. T. Caskey and Prof. Z. Chen

14. The Human Genome Project and clinical medicine: a progress report  
    V. McKusick (Baltimore, USA)
15. Esophageal cancer disease genes  
    M. Wu (Beijing, China)

16. Genes causing congenital malformations and heritable cancer  
    syndromes: The example of the RET proto-oncogene  
    G. Romeo (Geneva, Italy)
17. Molecular genetics of cystic fibrosis and other diseases associated  
    with human chromosome 7  
    L.-C. Tsui (Toronto, Canada)

**Break**
18. Positional cloning of X-linked disease genes  
A. Monaco (Oxford, UK)  
19. Genes related to human primary hepatic cancer  
J. -R. Gu (Shanghai, China)  
20. Genes involved in acute promyelocytic leukemia (APL)  
Z. Chen (Shanghai, China)  
21. Molecular genetic basis of amyotrophic lateral sclerosis  
H. X. Deng (Hunan, China)  

13:30 - 18:00  
**Session 3 Human Diversity**  
Chairmen: Dr. M. D’Urso and Prof. J. -Y. Chu  

22. Human genetic diversity and its evolutionary implication  
T. Gojobori (Mishima, Japan)  
23. Fourteen year’s studies on the genetic diversity of ethnic group in China  
R. -F. Du (Beijing, China)  
24. Twenty five ethnic minorities in Yunnan Province: their genotypic and phenotypic characteristics  
J. -Y. Chu (Yunnan, China)  

**Break**  

**Session 4 Other Genome Projects**  
Chairmen: Dr. G. Bernardi and Prof. J. -R. Gu  

25. The status of the Yeast Genome Program  
B. Dujon (Paris, France)  
26. Chinese Rice Genome Project and a progress of PKU lab  
Z. -L. Chen (Beijing, China)  
27. Rice molecular mapping and gene location  
L. -H. Zhu (Beijing, China)  
28. Analysis of some rice gene sequence and analysis of rice chromosomes  
C. -R. Sun (Shanghai, China)
29. Identification of DNA markers linked to blast resistance genes in rice  
   K.-L. Zheng (Beijing, China)
30. Sequencing, identification and mapping of *Schistosoma mansoni*  
   genes  
   S. Penna (Belo Horizonte, Brazil)

**November 9, 1994**  
**8:30 - 12:00**

**Session 5 Technical Advances**  
Chairmen: Dr. S. Penna and Prof. H.-X. Deng

31. Automated DNA sequencer with a new-continuous detection system  
   for simultaneous on-line sequencing of 44 clones  
   W. Ansorge (Heidelberg, Germany)
32. Overview of genome sequencing strategies  
   E. Chen (Foster City, USA)
33. New biology needs new approaches: DNA sequencing by hybridization and protein mapping on DNA by crosslinking  
   A. Mirzabekov (Moscow, Russian Federation)

**Break**

34. Construction of primers by ligation of short degenerated oligomers  
   and rapid sequencing without conventional cloning  
   W. Szymbalski (Madison, USA)
35. Fluorescent differential display: a high throughput system for a large-scale scanning of the differentially expressed genes  
   Y. Sakaki (Tokyo, Japan)
36. The application of molecular biology in Human Genome research  
   P. J. Ulfendahl (Pharmacia Biotech Ltd., Sweden)
37. New progress in technology of bio-engineering at Millipore  
   Y. -J. Liu (Millipore China Ltd.)
38. Silicon Graphics products introduction  
   J. -Y. Zhang (Silicon Graphics, Inc., USA)
39. New analyses and purification techniques in biology  
   H. -J. Liu (Waters Chromatography Division, USA)
Session 6 Genome Projects in Developing Countries
Chairmen: Dr. W Schlesinger and Prof. J. -H. Chai

40. Development of simple molecular biology techniques for genome research in Latin America
   S. Penna (Belo Horizonte, Brazil)
41. Research on human genetic diseases in Tunisia
   K. Dellagi (Tunis, Tunisia)
42. Studies on Theileria parva an intracellular protozoan parasite that provides a model of mammalian lymphocyte transformation
   O. Ole-Moi Yoi (Nairobi, Kenya)
43. YAC cloning and mapping of human X-chromosome and study of the DMD gene
   J. -H. Chai (Shanghai, China)

Break

16:00 Closing Ceremony
Chairman: Prof. S. Grisolia

44. New concepts in pathogenesis of slow viral infections
   Dr. C. Gajdusek (Bethesda, USA)
45. The Patient Recovers
   Dr. E. Dimond (Missouri, USA)
46. Medical genetic services in human genome research
   Dr. N. Fujiki (Fukui, Japan)
47 Mr. M -Z. Chen, Minister, Ministry of Public Health of China
SUMMARY OF PRESENTATIONS
Dear Director-General Dr. Mayor,

Distinguished guests, ladies and gentlemen,

We are very pleased to gather here to celebrate the opening of the “Second South-North Human Genome Conference” organized by UNESCO, the Chinese National Commission for UNESCO and Peking University. It is an honor for China’s Science Community to host for the first time an international conference in the genome research field. On behalf of the Chinese government and Science Community, I would like to extend my warmest welcome to all participants over the world, and to congratulate on the opening of this important event.

As a core part of the molecular biology and the frontier of modern sciences, the Human Genome Project was initiated in the mid-eighties. Since then, a great deal of achievements have been made. Many countries have joined the project and China is one of them. As early as 1989, Chinese scientists, having realized the extreme importance of the project, decided to take part in it. Following the idea of the genome project, and taking into account of circumstance of the country, China initiated Rice Genome Project in 1992. This project obtained the largest finding ever in history for biological research field from the State Commission of Science and Technology. Today, six Chinese laboratories have joined the project, the coordinators of this program are Prof. Guo-Fan Hong from the Institute of Biochemistry in Shanghai, and Prof. Zhang-Liang Chen from Peking University. Through their joint efforts, they have succeeded in constructing a BAC library, a RFLP map of rice, and have isolated a number of genes with agricultural importance. They have also achieved some progress on chromosome sorting. In addition, Chinese scientists are in the process of initiating genome programs for other organisms, such as wheat, corn and pig etc.
China is a giant agricultural country with long history of civilization. Paleoanthropological evidence shows that the Peking Man, flourishing in China some 400,000 or 500,000 years ago, knew how to make fire and use simple tools. Recorded history can be traced back to five thousand years ago. Now we have fifty six nationalities with a total population of 1.2 billion, and have formed a huge gene pool with great genetic diversities. It is incumbent for the scientists of China to share data resulting from the achievements of international Human Genome project, and be obliged to contribute to it by participating in this great international endeavor.

Having embarked along the way of reform and opening-up to the outside world, the Chinese government has payed great attention to the international scientific collaboration, including genome projects, the frontier of biology. I believe, this conference will catalyze active interaction between Chinese scientists and their international colleagues, and will stimulate the involvement of the developing countries in the human genome and other genome projects.

I wish the conference a resounding success and that all participants benefit from this experience.

Thank you!
Address by Ms. Wei Yu
Vice-Chairperson of the State Education Commission and Chairperson of the Chinese National Commission for UNESCO
at the Opening Ceremony of the Second South-North Human Genome Conference
(November 7, 1994)

Honorable Mr. Song Jian, the State Councillor,
Honorable Mr. Mayor, the Director-General of UNESCO.
Distinguished Participants,
Ladies and gentlemen,

Please allow me, first of all, to represent the State Education Commission and the Chinese National Commission for UNESCO and express my warmest congratulations to the opening of the Second South-North Human Genome conference and extend to all the participants my most cordial welcome.

I feel particularly Honored by the presence of Mr. Song Jian, the State Councillor and Chairman of the State Science and Technology commission who has come to address the meeting in person in spite of his tight schedule, and also of Mr. Mayor, the director-General of UNESCO, who traveled from a far to join us in the opening session today. This amply demonstrates the great importance attached by both the Chinese Government and UNESCO to this Conference.

To facilitate the meeting, UNESCO, the Academy of the Third World, the State Science and Technology Commission and the State Commission of Foundation for Natural Sciences have provided financial support, to whom I express my heartfelt thanks.

Fellow participants,
Ladies and gentlemen,

This is the first time for China to host a meeting in the field of human genome. Prestigious experts and scholars from all over the world are meeting here to conduct in-depth discussion on the research and development of human genome. As is known to everybody, life science is a newly-emerged discipline. As the most
conspicuous research programme within the life science. Human genome research programme is also an extensive systems engineering, calling for the joint efforts and cooperation among scientists worldwide. This conference is exactly the manifestation of such an international collaboration.

The implementation of the human genome engineering exerts significant impact on the development of life science research and biotechnologies. Therefore, both developed countries such as USA, Japan and France and some developing countries have set up their own research programs and centers so far. The Chinese Government is very much concerned with the development in the area. Quite a number of Chinese scientists are very active in this field. At this moment, they are preparing for the formulation of a human genome research programme in China and are seeking extensive international cooperation for the purpose. I am convinced that the convening of this conference will provide great impetus to the development of the research work in this field.

Mr. Chairman,

With the furthering unfolding of the reform and opening to the outside world, the scientific community in China is faced with unprecedented opportunities and challenges and the Chinese scientists are fully aware of their historical responsibilities. I believe that the joint efforts by scientists from China and the rest of the world will contribute significantly to the progress and development of humanity.

Mr. Chairman,

Last but not least, I would mention here that I am satisfied with the cooperation between UNESCO and China in the field of science. UNESCO has the capability to undertake joint research programs in a number of key or frontier disciplines with member states and China, together with other member states, would encourage UNESCO to do so. It is my sincere hope that the cooperation between UNESCO and China will be further strengthened.

I wish the conference a full success.
Thank you.
ADDRESS BY ZHANG CUN-HAO
Zhang, Cun-hao. Director. National Natural Science Foundation of China

Honorable Professors Grisolia and Chen!
Honorable Dr. Song Jian!
Honorable Dr. Federico Mayor!
Distinguished Guests from abroad!
Dear Colleagues!

On behalf of the National Natural Science Foundation of China, I would like to convey my warmest congratulations to the inauguration of the Second UNESCO South-North Human Genome Conference and at the same time my heartiest welcome to all our distinguished guests from abroad.

During the past few decades, not only has the science of genes been developed to an unprecedented maturity, but also has it become one of the most vital and incisive tool that almost all branches of life science are bound to rely on. It is gratifying to note that the human genome project is a spectacular outcome of such development in recent years. The growing importance of uncovering the human genome mystery has been recognized by the whole science community and has already greatly spurred ingenious research in numerous crucial fields.

Now let us recall that the 21st century is only 5 years ahead of us. It is certain that the 21st century will count even more heavily on life science and life-science-based technology than the present one. The critical role that the human genome project is going to play can hardly be overestimated. The science community may easily link it with the building of a much more splendid civilization that what we have today. So it appears a very meaningful move for the life scientists of all countries, whether north or south, to strive for the best in this respect.

In the past couple of years, the National Natural Science Foundation of China has already been supporting the human genome project, albeit in a relatively small scale. Spurred by this conference, I believe that NSFC would be willing to continue and, very hopefully, to extend its support on outstanding research programs devoted to this project, both within China and in international collaboration, I firmly trust that the present conference will make prominent contributions to this
crucial discipline in all countries represented and wish the conference every success
I also wish all participants from abroad a pleasant stay in China.

Thank you very much!
Hu, Zhao-guang, Vice Mayor of Beijing

Honorable Dr. Song Jian, Federico Mayor and Mr Chairman,
Ladies and Gentlemen:

Today the Second Human Genome Conference begins herein Beijing, China. On behalf of your host, the Beijing Municipal Government, I am very glad to give our warm welcome to you. Please receive my cordial congratulations. This conference provides our Chinese scientists in the field of genome research with an opportunity to exchange their experiences with the overseas experts. This will surely play a significant role on the accomplishment of our human genome project.

I wish you and the conference a complete success.

Thank you!
AN ADDRESS TO THE SECOND SOUTH-NORTH HUMAN GENOME CONFERENCE

Guan-hua Xu, Vice President, Chinese Academy of Sciences

Mr Chairman,
The State Councilor, Dr. Song Jian.
UNESCO Director-General, Dr. Federico Mayor,
Distinguished guests, ladies and gentlemen,

First of all, please allow me on behalf of Chinese Academy of Sciences to greet the opening of the Second South-North Human Genome Conference. The human genome research is very important and magnificent. The mapping of all genes and sequencing DNA of the whole genome of human being will make us understand the structure of all genes and know much more of their function and regulation. Moreover, we can explore much more easily then the genesis and development of human being and obtain important and useful information of potential and actual medical interest directed to human health and welfare. As we know that China has a population of about 1.2 billion, making up one-fifth of the world population. Besides, there are 56 ethnic groups. Therefore, we are extremely interested in exploring the characteristics of the Chinese genome. Just take our Academy for example, there are more than 10 research institutes and groups engaged in gene research. The Institute of Genetics of Chinese Academy of Sciences has taken part in the initial work of the Human Genome Diversity Project organized by US professors from Stanford University, Now the Chinese Committee of the Human Genome Diversity has been organized. The Institute of Genetics together with the Yunnan Medical Biology Institute and Harbin Medical University has started to establish relative data bank of Chinese population, Now, we are very pleased to learn that UNESCO has paid much attention to the human genome research and has made great contributions to the promotion of Human Genome Program in the world.

Today, we are extremely glad to see that Dr. Mayor himself and other UNESCO officials come here to attend this conference.

Dear friends, ladies and gentlemen,
In recognition of Dr. Mayor’s contributions to the advancement and promotion of science and technology as well as culture in the world, a great number of research and educational organizations in the world have awarded Dr. Mayor honorary titles. Now, it gives me a great honor on behalf of Chinese Academy of Sciences and ex-president, Professor Zhou Guangzhao to take this opportunity to confer the title of Guest Honorary Professor on your excellency Dr. Federico Mayor, Director-General of UNESCO.

Thank you!
I wish to thank each and all the organizers, but especially Professor Chen, for the efficient and generous reamer in which they have worked tirelessly to make this second UNESCO South North Conference in the Human Genome a success.

In an address by the Director General of UNESCO, Federico Mayor, to the conference held in Washington some 3 years ago under the title, “Human Genome Research in an Interdependent World”, Dr. Mayor clearly outlined the two main reasons which led to the development of the UNESCO Scientific Coordinating Program. These were:

1) The task of sharing new insights into the mechanisms of human heredity.

2) Ensuring that adequate thought is given to the wise application of the scientific knowledge obtained in this field.

UNESCO is entrusted by the world community with the promotion of international co-operation in science, education and culture and thus has a natural role in the human genome programme.

UNESCO’s interest in the human genome project started in 1988, by the support of, and participation in Valencia in the 1st meeting for International Cooperation. In the following year it invited some scientists to form a committee (the SCC) to guide UNESCO in the planning and implementation of its programme on the human genome.

On the recommendation of the SCC, UNESCO’s participation in the human genome project is on three main fronts: the co-ordination and integration of
international research efforts and dissemination of the results, the involvement of
developing countries, and stimulation of debate on the ethical, social, legal and
commercial aspects of the project. Because of the importance of Bioethics, Dr.
Mayor set up recently an independent International Bioethics Committee to deal
with these issues.

UNESCO has wide experience in international cooperation both with its
member states and non-governmental organizations. It provides a focal point for
exchange of data, technology and samples relevant to genome research and also for
debate between scientists from different disciplines and from widely separated
countries.

One of the most cost effective means of strengthening the scientific and
 technological capacities of less developed countries is through training and the
establishment of ties with laboratories in developed countries. UNESCO has helped
international conferences and workshops in the developing world on the human
genome and the techniques used for its analysis. Also in conduction with the Third
World Academy of Sciences (TWAS) UNESCO has funded short term fellowships
to allow students from developing countries to work in laboratories involved in
human genome research in developed countries. Thus far Fellowships have been
given to naturals from nations in the five continents.

A main criticism, and to some extent true of the Vaiencian and Russian
meetings has been that by and large these have been American workshops, with a
minimum of Oriental and European participation, while the developing world was
scarcely represented. Of necessity, this was largely true and this is why the
Scientific Coordinating Committee decided to exert a great deal of effort in
supporting a South-North main conference every year and, at the request of Dr.
Pena, Brazil was selected for the first, and we plan for next year to hold the 3rd
one in India.
In resume. I suggest that cooperation on the Genome Project at an international level although difficult, is necessary and reachable. depending on continuous and vigorous expositions and discussions such as the ones which will be carried out during this workshop.
The Chinese Human Genome Initiative
Boqin Qiang, MD
Institute of Basic Medical Science, Chinese Academy of Medical Sciences, Beijing

The Human Genome Project, which was pioneered by developed countries, is now in full swing and has achieved its most scheduled goals. Chinese biologists highly appreciate its foresight and fruitful progress. The Chinese, making one-fifth of the world population, are not only entitled to share data resulting from the achievements of international Human Genome Project, but also obligated to share the responsibility and effort by participation in this international project. The Chinese Scientists realize it is very important to have their own Human Genome Project with its own characteristics to face the challenge and opportunity. The National Natural Science Foundation of China (NSFC), the country’s main finding agency for basic research, launched a national human genome project in 1993 as part of its contribution to international efforts to sequence the human genome.

The Chinese Human Genome initiative is mainly aiming at: (1) Development and improvement of the techniques for genome research. (2) Collection and storage of genetic samples. (3) Analysis of some particular disease-related loci in Chinese genomes. In addition to the Han nationality, which is the largest gene pool in the world, China has more than 55 ethnic groups with a total population of 91.2 million (1990), or 8.04% of the whole population in China. The Chinese Genome Initiative Project is also designed to study genetic diversity of the different ethnic groups and to find out the frequencies of main alleles at specified loci, and differences between them. The scientific goals of the genetic investigation of ethnic groups are to determine the types and incidence of genetic diseases and to preserve the genetic materials.

As one of the developing countries, although hampered by financial problem, the Chinese Human Genome Project has entered its beginning stage. A YAC library has been introduced from CEPH, which has shown interest in collaboration with the
Chinese Genome Project. A YAC library containing the entire genome of a Chinese individual is under construction. Cancer-related genes in several cancers, such as esophageal cancer and liver cancer have been localized or cloned, Progress has also been made in several chromosomal regions, e.g. the DMD region, region around the fragile X, and chromosome 17. New techniques have been developed. Attention has also been paid to related issues, such as training, education, ethical, legal and social implications of the genome project.

The Chinese Genome Initiative will be totally open to the world and expect desperately the understanding, acknowledgment, moral, financial and material support from the International Community. The future will also witness that the International community will benefit from Chinese participation, from its unique and largest biomaterial resources.

The organizations for CHGP
1. Academy committee:
   QIANG, Boqin Coordinator
   M. D., Professor,
   Member of Chinese Academy of Sciences
   Vice President of Chinese Academy of Medical Sciences
   Peking Union Medical College
   5 Dong Dan San Tiao, Beijing 100005

   CHEN, Zhu Coordinator
   Ph. D., Professor
   Director of Shanghai Center of Human Genome Research
   Shanghai 2nd Medical University
   197 Rui Jin Er Lu, Shanghai 200025

   CHEN, Renbiao Ph. D. Professor
   Shanghai 2nd Medical University
DENG, Hanxiang  Ph. D., Professor  
Director of National Lab of Medical Genetics. Hunan Medical University  

DU, Ruofu  Ph. D. Professor  
Institute of genetics, CAS  

SHEN, Yan  Associate Professor  
Institute of Basic Medical Sciences, CAMS  

CHU, Jiayou  Ph. D. Professor  
institute of Medical Biology, CAMS  

Secretary-General:  
YANG, Huanming  Ph. D. Professor,  
National Center of Genetical Medicine CAMS  

2. National Advisory Committee  
TAN, Chiachen  Honorary Advisor, Ph. D. Professor  
Member of the Chinese Academy of Science  
Fudan University  

WU, Min  Chief Advisor, M. D., Professor  
Member of the Chinese Academy of Science  
Chairman of Dept. of Life Sciences, NSFC  
Director of National Lab of Molecular Oncology  
Institute of Cancer, CAMS  

SHI, Luji  Vice-Chief Advisor, Ph. D., Professor  
Member of the Chinese Academy of Science  
Institute of cell Biology, CAS  

HONG, Mengming  Vice-Chief Advisor, Ph. D., Professor
Member of the Chinese Academy of Science
Institute of Plant-physiology, CAS

CHAI. Jianhua  Ph. D. Professor
Institute of Genetics, Fudan University Shanghai

GU, Jianren  M. D. Professor
Director of National Lab of Oncogenes & related Genes.
Shanghai Institute of Cancer Research

LI, Pu  Ph. D. Professor
Harbin Medical University

LO, Huiyuan  M. D. Professor
Director of National Center of Genetic Medicine,
Beijing, CAMS

QI, Shuying  Ph. D. Professor
Vice Chairwoman of Dept. of Life Sciences, NSFC

WU, Guanyun  Ph. D. Professor
Institute of Basic Medical Sciences CAMS

ZHAO, Shouyuan  Ph. D. Professor
Institute of Genetics, Fudan University

3. Labs Involved in CHGP
(l). National Lab of Medical Molecular Biology
   Institute of Basic Medical Science (IBMS)
   Chinese Academy of Medical Science (CAMS)
   Beijing
(2). National Lab of Molecular Oncology
    Institute of Cancer Research CAMS, Beijing

(3). National Center of Genetical Medicine
    IBMS, CAMS, Beijing

(4). Institute of Medical Biology
    CAMS, Kunming, Yunnan

(5). Institute of Genetics
    Chinese Academy of Science (CAS)
    Beijing

(6). Institute of Biophysics, CAS, Beijing

(7). Institute of Zoology, CAS, Kunming, Yunnan

(8). Dept. of Medical Genetics
    Harbin Medical University
    Harbin, Heilongjiang

(9). Shanghai Center of Human Genome Research
    Ministry of Public Health, China
    Shanghai 2nd Medical University, Shanghai

(10). Dept. of Medical Genetics
      School of Basic Medicine
      Shanghai 2nd Medical University, Shanghai

(11). National Lab of Life Sciences
      Institute of Genetics
      Fudan University, Shanghai
(12). Institute of Cell Biology, CAS. Shanghai

(13). Shanghai Institute of Medical Genetics

(14). National Lab of Oncogenes and Related Genes  
    Shanghai Institute of Cancer Research

(15). National Lab of Medical Genetics  
    Hunan Medical University  
    Changsha. Hunan
THE DISTRIBUTION OF GENES IN THE HUMAN GENOME
G. Bernardi.
Laboratoire de Genetique Moleculaire, Institut Jacques Monod,
2 Place Jussieu. 75005 Paris, France

The first part of this presentation will briefly review results from our laboratory on the isochore organization and the gene distribution of the human genome, the second will discuss their implications for genome projects and other related issues.

**Isochores.** The human genome is a mosaic of isochores, long (>300 Kb) DNA segments, which are compositionally homogeneous and can be subdivided into a small number of families characterized by different GC levels covering a 30-60% range (1-3, GC is the molar ratio of guanine + cytosine). Isochore families L1+L2 and H1+H2+H3 represent the GC-poor 2/3 and the GC-rich 1/3 of the genome, respectively. The isochore organization of the human genome is typical of warm-blooded vertebrates, whereas cold-blooded vertebrates are characterized by a much narrower GC range which never reaches the high levels attained by warm-blooded vertebrates. The different isochore patterns just mentioned are paralleled by different compositional patterns of coding sequences. The different compositional patterns of isochores and coding sequences of warm- and cold-blooded vertebrates represent different genome phenotypes.

**Isochores and chromosomes.** *In situ* suppression hybridization of human DNA fractions characterized by increasing GC levels on human metaphase chromosomes has clarified the correlations between DNA base composition and chromosomal bands (4,5) : (i) T(elomeric)-bands are formed by the GC-richest isochores of the H3 family, and by part of the GC-rich isochores of the H1 and H2 families (with a predominance of the latter); (ii) R'-bands (namely R(everse)-bands exclusive of T-bands) are formed, to almost equal extents, by GC-rich isochores of
the H1 families (with a minor contribution of the H2 and H3 families) and by GC-poor isochores of the L1+L2 families: (iii) G(iemsa)-bands essentially consist of GC-poor isochores from the L1+L2 families. with a minor contribution of H1 isochores.

**Gene distribution.** The distribution of genes in the human genome is strikingly non-uniform (1-3). Indeed, a low, constant gene concentration is present in the GC-poor isochores L1 and L2; gene concentration then increases in increasingly GC-rich isochores (isochores families H1 and H2) to attain the highest value (20x higher than in L1+L2) in the GC-richest isochores family H3, that only represents about 4% of the genome. Because isochores distribution in chromosomes is known, these results also provide information on the distribution of genes in chromosomes. The highest concentration of genes in telomeres is of special interest in view of the association of telomeres with the nuclear membrane and of their attachment to the nuclear matrix.

The correlation between gene concentration and GC level stresses the interest of compositional mapping, namely of establishing an isochores map in physically mapped regions of the genome, since an isochores map is a gene concentration map.

The reasons for the correlation between gene concentration of isochores and their GC levels have been recently understood. Indeed, the gene concentration pattern of the human genome is basically present in all vertebrates. A strong GC increase took place, however, in gene-rich regions at the transition between cold-blooded vertebrates on the one hand and mammals and birds (two separate events in time) on the other. A comparison of aligned homologous coding sequences unequivocally demonstrated that GC changes were caused by directional fixation of mutations, which may be due to a mutational bias and/or to natural selection.

**Functional significance of isochores.** Several findings point to the fact that isochores have not only a structural but also a functional relevance. For
example, the GC-richest isochore not only are characterized by the highest gene concentration, but also by the highest concentration of CpG islands (GC-rich, non-methylated, regulatory regions located upstream of genes), by an open chromatin structure, by the highest transcription and recombination levels, and by an extreme codon usage and amino acid encoding.

**Implications.** The relevance of the notions outlined above for genome projects and other related issues will be illustrated by discussing: (i) the causes of chromosome banding (6); (ii) the generality of the isochrome organization in eukaryotes (7-10); (iii) the generality of the correlation between gene density and GC, as exemplified by investigations on yeast chromosome IX (10); (iv) the coverage of genes by the first generation physical map of the human genome (11); (v) the use of H3 isochore for completing the human genetic map (12); (vi) the sequence conservation between the murid and the human genomes (13); (vii) the proposal of puffer fish as a model system for the human genome (14); (viii) isochore and the long-range correlations in DNA (15); (ix) the isochrome-specific integration of viral sequences (16); and (x) the very striking gene distribution of the maize genome (Carels et al., paper in preparation).

MAP-DRIVEN GENE FINDING TO FULFILL THE GENOME PROJECT

David Schlesinger
Washington University School of Medicine, St. Louis, MO 63110 USA

From its inception the major driving force for the Genome Project has been the medical interest in identifying disease genes and analyzing the pathophysiology that they produce. But the scientific community has also realized that once the gene complement and corresponding sequences of the genome are known, they become the starting point for studies of human biology for the rest of time.

The justification of the Genome Project is that it would provide economy of scale and new technology so that studies of genetic diversity, genetic morbid anatomy, evolution, and chromosome dynamics could proceed faster and at lower cost. It has been estimated that 80% of the cost of disease gene searches in the 1980’s was involved in obtaining DNA to cover a region that had been found to harbor a disease gene by cytogenetic or linkage mapping studies. Thus, if the entire genome could be assembled in overlapping DNA clones -- and even better, if a catalogue of all the genes in the corresponding region could be known to the sequence level -- no one would have to reclone the DNA for any region de novo again; and 80% or more of the cost of a gene search would be cancelled.

To underline the potential power of this approach, an example is provided by the cystic fibrosis gene. In that case, the groups of Drs. L. C. Tsui and F. Collins that were instrumental in the landmark discovery of the gene were among 50 laboratories, with an estimated expenditure of $50 to $100,000,000 during the search. Few of the 100,000 genes in the genome would warrant that kind of level of expenditure, and no research program could afford to find a large fraction of genes with that cost schedule. On the other hand, starting from the same linkage information and a limited number of markers in the region, it was possible for Eric Green to assemble the same region in cloned DNA in a short time and at low cost.
containing the entire gene intact for the first time, using the YAC cloning systems that had been developed for mapping in the human genome project.

Today no one disputes the utility and feasibility of the achievement of the physical map and its integration with markers that detect polymorphism. To derive the maximum benefit from the resultant map, however, the problem of scientific interest and the technological challenge now shifts to how best to use the mapped DNA to drive gene searches.

We start from the formulation that the map to be used is provided by contigs of large-insert DNA clones, currently yeast artificial chromosomes (YACs) for the most part, and that these are formatted with the sequence tagged sites (STSs) defined by PCR primer pairs. How are we to find a disease gene— or all the candidate genes — in a particular region?

The aid which is provided by YACs is based on two features, both related to their size: 1) they make long-range coverage of a region simpler; and 2) they are large enough to contain even a sizable gene intact and in normal regulatory context.

Based on these starting materials, a number of approaches have been successfully used to find genes. Here we detail a number of examples.

At present one can define two extreme paradigms for gene-finding. "paradigm 1" starts by detecting and analyzing cDNAs. It then localizes the cDNAs on corresponding genomic DNA. Some of the growing number of methods employed are summarized below.

The alternative "paradigm 2" works in the opposite sense. It begins with systematic sequencing of genomic DNA and then predicts possible exons in the sequence. The existence of the putative genes is then assessed by searches for corresponding cDNAs. Such an approach has in fact worked for several DNA and RNA viruses, and the yeast and nematode projects have already made sequencing
and computer-assisted gene searches cost-effective for many areas at the megabase level. At present, mixed paradigms are in use. Here we review the prospects for gene-finding by long-range physical mapping and sequencing compared to cDNA or motif localization. Here I review some of the techniques; in the presentation, further examples will be given for various techniques, using genes from the X chromosome.

**Paradigm 1. Functional mapping by cDNA identification**

**Localization of cDNA sequences.** This approach has often been reviewed. It identifies genes by the direct study of RNA transcripts and cloned DNA copies made by reverse transcription. An extreme variant is the collection of bits of sequence from large numbers of individual cDNAs in commercial and governmental projects. An advantage of this approach is its rapid and high yield of tags for genes. The disadvantages include the dependence on collections of cDNAs which vary in quality and representation; the variable fidelity of the short tracts determined in one pass sequencing, which can make it particularly difficult to discriminate the many transcripts belonging to a gene family; and the tendency of various projects to repeat the identification of the same RNA species.

The direct recovery of cDNA species that hybridize to a particular genomic clone is a powerful variant of localization. The probes and methods to search for particular species in cDNA collections now include a number of methods based on the use of genomic DNA, for example in recombination-based and direct selection methods.

**Identification of genes by telltales.** A large range of sequence elements or defined sequences that are not unique can nevertheless be used to locate possible genes. A frequent approach looks for clusters of rare-cutter restriction sites that
include CpG dinucleotide sequences: such "CpG" islands are invariably associated with genes, and are found in promoter or exonic regions of about 1/3 of all genes.

A generalized form of searches for "islands" is provided by determinations of GC level and CpG dinucleotide enrichment across genomic regions. Overall GC content is an index of gene content in subchromosomal regions. The highest GC fraction is highly enriched for genes, and falls in 3 to 5% of total human DNA at so-called "T" cytogenetic bands, mostly in subtelomeric regions.

A group of moderately repetitive sequences more specific than CpG content and especially interesting in its own right encode portions of proteins that are conserved during evolution. These BLOCKS have been analyzed in increasing detail, and on the order of 2000 BLOCKS have been identified. The approach is likely to be useful because up to 40% of genes are thought to be associated with at least one BLOCK.

Telltales for genes can be functional as well as structural. They of course include both the linkage mapping approach that characterizes the positional cloning of disease genes and genes specifying other traits. In recent approaches, fictional complementation can be used to recover the genomic form of genes, for example for a tumor suppressor.

Paradigm 2. Functional mapping based on sequencing.

Moving toward sequencing from contigs of clones is not routine or even straightforward. Nevertheless, it is already possible to sequence intervals of hundreds of kilobases, with increasing efficiency.

At present, a number of computer-aided searches are available to hunt for possible genes and other functional elements in a sequence. For example, a cosmic centromeric of the glucose 6-phosphate dehydrogenate (G6PD) gene in Xq28 was completely sequenced. Searches for total GC revealed two spikes of very high GC
DNA (80%) with several rare-cutter sites suggestive of CpG islands: one of these was already known to be at the 5' end of G6PD. Searches with gene-hunting programs then detected three additional sites at which genes were further suggested by the locations of clusters of transcriptional motifs. Thus far, for two of three predicted genes, primer pairs bracketing two putative exons were synthesized, and PCR indeed confirmed the presence of corresponding cDNA species.

PROSPECTS

Searches for predicted exons and searches in sequence for genes mapped to a region have already been successful in pilot cases. At present, in model organisms like E. coli, yeast, and nematode, the relative simplicity of genome structure has facilitated the finding of genes with a net yield in nematode DNA of about 1 gene per 7 kb of sequenced DNA. The concentration of genes in the GC-richest fraction of human DNA is comparable to that in the nematode, making that fraction a prime substrate for long-range sequencing with current technology. A three- to 10-fold further increase in sequencing efficiency would tip the balance further toward large-scale sequencing as a realistic step in an approach to the complete catalog of human genes.
As the indispensable first step in genome analysis, technological barriers were overcome, and successful mapping of chromosomes in overlapping clones is progressing. The map is the precondition to the further development and fruition of the project goals of developing a catalogue of genes for the entire genome. A widely accepted feature of the map is the development of ordered "sequence-tagged sites" or "STSs" as landmarks. Each STS is defined by an oligonucleotide primer pair that produces a unique PCR product from genomic DNA. With the human genome of $3 \times 10^9$ basepairs and an estimated 100,000 genes, an STS every 100kb would provide a marker stationed on average near every second or third gene. An "average" 150 Mb chromosome would thus require on the order of 1500 STSs.

Most maps are currently constructed from overlapping yeast artificial chromosomes (YACs). The first large-scale application of YAC technology to a segment of the human chromosome, and one of the furthest toward a high-resolution map, was the analysis of the portion of the X chromosome that is reviewed here. Special features of the X chromosome rendered it a prime choice for increasingly intensive analysis, both for fundamental research and for its practical applications. First, the X chromosome, because of the mechanism of sex determination in mammals, was separated off from the other chromosomes early on in evolution, and its content of genes became for the most part fixed, The biological consequences of this isolation of the X are profound and various, including the extraordinary phenomenon of X chromosome inactivation. The analysis of the X is also a first step toward the study of the evolution of the sex chromosomes, and can be consistently compared to the map of the Y chromosome. Second, the ease of genetic assignment has meant that data began to accumulate about the X chromosome early on. As a result, there are many traits, markers, probes, etc.
available, making the result more immediately useful for many practitioners of human genetics. Third, there is a very large amount of observed X-linked human pathology. For example, the Fragile X syndrome, genetic analysis has inferred that there are more than 15 genes on the X defects in which lead to mental defectiveness. X-linked maladies are both easier to observe and more likely, since the single X in males exposes recessive mutations,

We chose to concentrate on 50 Mb at the distal end of the long arm of the X chromosome-- a region that was known to contain important genes involved in disease, including both the Factor VIII and Factor IX genes involved in the etiology of hemophilia, the HPRT gene involved in Lesch-Nyhan syndrome, the Fragile X locus, and above all a curious concentration of disease genes that linkage mapping had begun to pile up in distal Xq28. The notion was to attempt to get reliable coverage of the region and to try to understand the differential accumulation of disease genes in part of the region,

The project began with the production of a targeted library of YAC clones for Xq26-qter. This was made by starting from the DNA of a hamster/human hybrid cell that contained the distal portion of the X chromosome as its only content of human DNA. Using these clones and supplementing them later with additional YACs from other collections, a frame work map covering much of the region was developed, overlapping YACs by their common content of hybridization probes. More recently, these efforts were followed up by the increasing development and mapping of STSs. These have progressively extended the YAC contigs and merged and formatted them,

YAC-based contigs now cover the region almost entirely, assembled and analyzed by their content of STS content, hybridization probes, and fingerprinting methods.

The longest set of unambiguously ordered and oriented contigs have been assembled across more than 31Mb of YAC contigs in Xq26-qter.
As in Xq26-qter, the region of Xq24-q26.1 has also been assembled in YAC contigs of average size greater than 2Mb. The region still contains a number of "zaps" where no coverage has been found in several libraries, but contigs have been further rationalized and ordering has been accomplished by the use of the linkage map and fluorescent in situ hybridization methods.

To warrant its use in fictional analysis, a collection of overlapping clones (a "contig") should be faithful to the structure of uncloned genomic DNA. At present it seems that the bulk of the genome is "well-behaved" in YACs. As an indication of the quality of the map, for example, the region of Xq26-q27.3 containing 21Mb of DNA in a single contig (Zucchi et al., ms. in preparation) includes more than 500 YACs, and overlapping YACs typically show consistent contents of 500 STSs and more than 800 hybridization probes.

Physical mapping is thus now reaching the phase of "closure". The last 20% of the map can, however, require considerable effort, with contigs presenting research problems of an idiosyncratic kind. Difficulties that have been encountered include the need to deal with regions that are unstable, recombined, or unrecovered in cloned DNA.

The Xq28 region, where my own group has worked closely with that of Dr. Schlessinger to develop a detailed map, shows both the degree to which a map can be constructed with current materials and exemplifies the remaining problems to be overcome. The region is covered in 5 contigs of YACs that include 7.1Mb of DNA; and a detailed restriction map has been produced that indicates the localization of rare-cutter restriction sites across the region.

This map is clearly adequate to support activities like gene-finding and the approach to sequencing. Most important, the apparent concentration of disease genes in the region of distal Xq28 can be explained as a consequence of the differential distribution of GC in subregions of the genome. Over Xq28, GC content
ranged from, 42% to 55%. As anticipated from Bernardi’s earlier studies. "isochores” of relatively constant GC level were observed, even across megabase stretches of DNA. Moreover, there was a correlation of GC level, CpG dinucleotide content (assessed by rare-cutter sites), and the content of the CpG islands that are found associated with many genes. The highest GC levels were observed in the subtelomeric region around the RCP/GCP locus, with up to 1 CpG island per 30-50kb. This is the same region in which disease genes had been “accumulating”, and it is now clear that more disease genes are located there because the region contains relatively more genes.

Ironically’, the finding from mapping that the GC-rich regions of high gene content are located, as in this case, subtelomERICALLY has been accompanied by the realization that those regions are selectively poorly recovered in YACs. In our studies, in the general experience of Centers doing large-scale screening, and in the published first-generation genome-wide efforts of the CEPH/Genethon consortium, it is precisely the high GC portion of the genome that is recovered poorly in YACs, and often in clones that show rearrangements.

The color vision region is typical of problems in such areas. Only small YACs are seen in a 1-1.5 Mb region from LICAM to G6PD, and sequence analysis of several of them demonstrates that they frequently show large deletions and possible rearrangements.

Thus, a major challenge of closure efforts is to refine the map and fill in portions that are not covered. Fortunately, there are good prospects for success, for several reasons, First, the map is complete enough to delimit the discrete number of sites in the genome that contain the highest GC DNA -- the 3% of “H3 DNA” with an estimated 1/4 or more of the genes, Second, much of these regions can be covered well in bacterial clones like cosmids and lambda phage. Third, some groups are producing YAC clones that show less tendency to rearrange and better retention of at least some high GC segments of chromosomal DNA; and even long-range PCR may develop to permit direct recovery of recalcitrant portions of the genome.
without any intervening cloning steps at all. Thus the remaining minor fraction of the Xq24-qter test region should also soon be on hand, and the next phase of sequence analysis is starting in full force.
ESOPHAGEAL CANCER DISEASE GENES

Min Wu
Cancer Institute, Chinese Academy of Medical Sciences.
Peking Union Medical College, Beijing 100021, China

Esophageal cancer is the second most common cause of cancer death in China and is particularly prevalent in northern China (1-3). The current view on the etiology of esophageal cancer largely favors environmental factors, and is supported by the marked variation in the geographic distribution of the disease. However, some recent studies have suggested that genetic factors may also play an important role in the development of this human malignancy (3-5).

Evaluation of a positive family history comes from epidemiologic studies of esophageal cancer in Shanxi, a high-incidence esophageal cancer province in North China. It turned out that family history was associated with a 4-8 - fold increase in risk for esophageal cancer (6, 7). A similar study in Linxian, another high risk county in Henan province, has also demonstrated increased risk among persons with a family history of esophageal / stomach cancer (8).

Using logistic regression models, segregation analysis was performed on 221 high-risk nuclear families from a village in Linxian who had at least one affected family member and all offspring over the age of 40 (9) This analysis indicated that there is a Mendelian pattern of transmission, most likely from an autosomal recessive gene with a frequency of 19%. Segregation analysis was also performed in another set of 284 nuclear families from high-risk pedigrees in Shanxi Province. The results again suggested a Mendelian transmission pattern, although the data here were most compatible with an autosomal dominant gene present in 5% of the population (unpublished data). Cytogenetic studies revealed that esophageal cancer patients and two-thirds of their blood relatives from high-risk families showed increased chromosomal instability (10 - 12),
On basis of the results obtained from cytogenetic and molecular studies, as well as the family data collected from high-risk areas studies on characterization and isolation of the gene(s) deleted frequently in esophageal cancer or germline of high-risk families are under way in our lab (13-15 and unpublished data).

In order to explore the new possibilities in therapy, a strategy for isolating differentiation-inducing cDNAs from human esophageal cancer cell lines treated with all-trans retinoic acid was developed in our lab and a series of cDNAs had been isolated (16). A cDNA of 3.8kb showed very strong biologic effects on human cancer cell lines. When this cDNA was transferred into and expressed in cancer cell lines the latter stopped to grow and underwent terminal differentiation and apoptosis (17,18).

References


16. FENG L, WANG XQ, FU M, WANG ZH, TIAN Y, CAI Y and WU M: A strategy for isolating differentiation-inducing complementary DNAs from...


GENES CAUSING CONGENITAL MALFORMATIONS AND HERITABLE CANCER SYNDROMES: THE EXAMPLE OF THE RET Protooncogene
Giovanni Romeo - Instituto Giannina Gaslini 16148 Geneva

The concept of phenotypic diversity (PD) and allelic series (AS) is supported by many examples arising from the direct identification of mutations in genes of patients with apparently unrelated disorders (1). An interesting recent exemplification of this concept is provided by the mutations of the RET protooncogene located in 10q11.2.

Germ-line mutations in the human RET protooncogene are involved in the etiology of 4 different inherited neural crest disorders (neurocristopathies): Multiple Endocrine Neoplasia 2A or MEN 2A (2, 3), MEN 2B (4), Familial Medullary Thyroid Carcinoma (FMTC) and finally Hirschsprung disease or HSCR (5, 6). HSCR is a developmental disorder of the enteric nervous system, inherited as an autosomal dominant trait with incomplete penetrance and variable expressivity. It is caused by mutations in different domains of RET. Mutations affecting specific cysteine residues of the extracellular domain of RET in exons 10 or J1 cause MEN 2A, while a specific mutation in exon 16 of the tyrosine kinase domain has been found in 95% of patients with MEN 2B. MEN 2B is phenotypically similar to MEN 2A but shares with HSCR the involvement of the enteric nervous system, represented by diffuse ganglioneuromatosis along the intestinal tract. From the structural point of view RET is classified as a tyrosine kinase receptor whose extracellular domain is in part homologous to that of cadherins (7, 8), a class of proteins responsible for homophilic cellular adhesion (9). In general tyrosine kinase receptors are involved in the activation and control of cellular proliferation (10). RET is made of 20 exons as recently established by our group which reported also the extended intronic sequences flanking each exon (11). These results are somewhat at variance with those of Kwock et al. (12). The explanation for this
major discrepancy (namely 21 exons reported by the latter group instead of 20) may be explained as an artifact caused by the exon-trapping strategy used by Kwock et al (12). The complete genomic map of the RET region, including its promoter, has been recently delineated by our group together with the orientation of the gene on 10q (13). This map is based on the restriction with 10 different endonucleases of a cosmid contig of the RET region.

In addition to RET there are at least 7 other oncogenes or tumor suppressor genes whose mutations may cause developmental disorders. These genes may therefore be classified also as "teratogenes". The genotype-phenotype correlations between mutations observed in these teratogenes and the corresponding malformations on one side and tumor phenotypes on the other should provide valuable insights regarding the functioning of these genes during development and oncogenesis.

The reason for the existence of diverse phenotypes arising from mutations in the same gene might only in part be due to the different domains involved by the mutations. Possible further explanations for phenotypic diversity can be: a) additional mutations in the same gene that modify the disease phenotype; b) interaction with mutations in other genes (modifiers). The study of phenotypic diversity may also yield clues to the understanding of incomplete penetrance and/or variable expressivity.

References

MOLECULAR GENETICS OF CYSTIC FIBROSIS AND OTHER DISEASES ASSOCIATED WITH HUMAN CHROMOSOME 7

Lap-Chee Tsui.
Department of Genetics. The Hospital for Sick Children and Department of Molecular and Medical Genetics, The University of Toronto. Ontario.
Canada M5G 1X8

Cystic fibrosis (CF) is an autosomal recessive disorder prevalent in the Caucasian population. Hallmarks of CF include highly elevated concentrations of sodium chloride in sweat, chronic obstructive lung disease and bacterial colonization, blockage of pancreatic ducts leading to pancreatic enzyme insufficiency and reduced fertility. The identification of the cystic fibrosis transmembrane conductance regulator (CFTR) gene was based on a genetic approach with which the position of the gene was localized to the long arm of chromosome 7 by family linkage analysis followed by molecular cloning and characterization of a large region of the genome for candidate genes. This method, now commonly known as positional cloning, is widely used in the study of many other genetic disorders, including some forms of cancers.

In the case of CF, the gene identification has led to our present understanding that CF is due to defect(s) in a cAMP-regulated chloride channel and provided an explanation for the observed clinical symptoms. The major mutation ΔF508 accounts for about 70% of all CF chromosomes worldwide and more than 450 mutations have been detected in CF patients by members of the CF Genetic Analysis Consortium.

Although the varied presentation of disease among different patients cannot be entirely explained by the heterogeneity of CFTR gene defects, there is good correlation between different classes of CFTR mutations and exocrine pancreatic function. While most of the common CFTR mutations correspond to severe loss of
function, at least 15 mild alleles belong to the class of mild mutations. It has also been possible to classify different mutations at the functional level of the CFTR protein with respect to chloride regulation in the apical membrane. The spectrum of mutations has therefore contributed to our knowledge of the structure and function of the CFTR channel. It is of interest to note that most of the mild mutations are amino acid substitutions within or near the transmembrane domains. The definition of these mild mutations has resulted in the observation that mutant proteins such as R117H only affected the channel properties but retained the cAMP regulation. More recently, the work in our laboratory has defined a group of mild mutations caused by aberrant splicing. There are also mild mutations found within the first ATP-binding domain and R-domain. The identification of CFTR mutations in atypical CF patients and other related diseases, such as absence of vas deferens and chronic obstructive lung disease, further suggests that clinical manifestation of CF is complex.

Although gene therapy appears to be an attractive solution to the lung disease in CF, detailed description of the molecular consequence of each mutation may offer other therapeutic strategies based on the fictional properties of CFTR. In addition, it is important to understand the role of CFTR in different tissues at the cellular and subcellular levels. Further, we need to identify other genetic and environmental factors that contribute to the clinical variability in CF.

To approach the latter issues, we have exploited the use of a mouse model by replacing the mouse CFTR protein with human CFTR. The strategy is to introduce the entire human CFTR cDNA into exon 1 of the homologous mouse gene through embryonic stem (ES) cells. The resulting animal would not only confirm functional equivalence of the human and mouse CFTR proteins, but would also allow introduction of mutations into the human protein in a single step. ES cells with appropriate insertion of wild type human CFTR and two mutant cDNA (ΔF508 and R117H) have been generated. In preliminary studies, mice with simple insertion mutation in exon 1 were established. The general phenotype of the mice with the exon 1 "nock-out" is similar to the other CF mice reported in the
literature (through disruption of other regions of the gene). Most homozygous CF mice died of intestinal obstruction soon after birth. Or at the time of weaning. The survival of mice also appeared to be strain dependent. These mice thus provide a novel means to identify the additional genetic factor(s) that may contribute to the CF pathophysiology.

In order to study the many other disease genes yet to be identified, it becomes apparent that a comprehensive description of the human genome will greatly accelerate the research. The availability of a complete human gene map will facilitate the search of genes for disease with reduced penetrance, variable expressivity, or poor diagnostic criteria. A gene map also provides a starting point for analyzing complex disorders, such as diabetes and arterial hypertension and common diseases such as neuropsychiatric disorders, cancer, growth abnormalities, and heart disease. In addition, a human gene map will provide new knowledge in understanding structure, organization and regulation of genes and DNA sequences important for normal cellular and body function. Thus, many large-scaled and organized studies have taken place in the past several years; the effort is often chromosome-based, with national funding agencies and international programs facilitating and coordinating data collection and exchange.

Our experience on the study of human chromosome 7 may be used as an example of genome research. The work is supported with funds from the Canadian Genome Analysis and Technology Program. The principal strategy of our project is based on the use of yeast artificial chromosome (YAC) technology. A 5x library has been constructed with over 2,000 human chromosome 7-specific clones of an average size of 500 kb. In addition, a set of reduced somatic cell hybrids have been collected for regional localization of gene and DNA segments to more than 25 cytogenetic intervals for the long arm of this chromosome. The YAC clones are being assembled into overlapping fragments by screening with known genes, previously isolated DNA segments, genetic markers as well as ends of YAC clones. Mapping of over 800 YAC clones has been achieved by hybridization with these probes as well as by fluorescence in situ hybridization (FISH). Clusters of
overlapping YAC clones (contigs) have also been constructed. Moreover, using direct cDNA selection and exon amplification procedures, we are also isolating gene directly with the YAC and cosmid clones. Therefore, our map consists of physical genetic, as well as coding information. The map has already allowed us to embark on many genes that are important for regulation of growth and development.

Together with Dr. James Evans (now at University of North Carolina at Chapel Hill), we have defined a set of cytogenetic breakpoints specific in patients with ectrodactyly (also known as split hand and split foot deformity). Physical mapping studies have localized these breakpoints to a small region at 7q21-22. Several candidate genes have been isolated. Two of them belong to a gene family related to the Drosophila Distal-less. Another which appears to encode a small acidic polypeptide shows localized expression at the limb bud. RNA in situ hybridization data are suggestive of strong expression at the interdigital mesenchyme regions of developing mouse limbs when the digits are formed. In addition, we are collaborating with other researcher groups on the study of other disease genes located on the long arm of chromosome 7, including holoprosencephaly, triphalangeal thumb, William syndrome, Smith-Lemli-Opitz syndrome, non-insulin dependent diabetes mellitus, obesity and leiomyoma.
Research in the Human Genetics Laboratory is directed towards understanding the molecular nature of genes responsible for human inherited disorders and cancers. Various genetic and molecular techniques are being used to isolate large DNA fragments containing genes on the human X chromosome and selected autosomal regions as part of the human genome project. Genes are being isolated by positional cloning techniques to identify mutations giving rise to disease and the relationship of genotype to phenotype. Ultimately, our goal is to develop better diagnostic tools and potential therapies by studying the protein products of these genes as well as their biological function and role in the pathophysiology of the disease.

Background

Positional cloning is the process whereby genes are identified as a direct result of genetic analysis and does not involve fictional information such as protein sequence or available antibodies (for review see Collins 1992). Positional cloning is a multistep process which begins by localizing a disease gene locus to a particular region of a chromosome by the identification of structural abnormalities and/or genetic linkage analysis in families segregating the disease. This preliminary localization is followed by a molecular analysis of the region including finer genetic mapping, physical mapping, DNA isolation, transcript identification, cDNA cloning and searching for mutations in candidate genes.

Most disease genes isolated to date by positional cloning techniques have been pinpointed by chromosome abnormalities that are cytogenetically visible such
as fragile sites, deletions, duplications and transactions, Physical mapping and DNA isolation of large regions surrounding the chromosome position of these abnormalities has been greatly advanced by the ability to clone large fragments of DNA as yeast artificial chromosomes (YACs, Burke et al., 1987). For the human genome project, YACs have been essential for the construction of physical maps of cloned DNA (Cohen et al., 1993) and have also been used to identify the position of structural rearrangements in patients’ chromosomes by fluorescence in situ hybridisation (FISH) of YAC DNA to metaphase chromosome spreads.

When there is no obvious structural aberration detectable cytogenetically or as altered DNA fragments on Southern blots, disease genes must be localized solely by genetic linkage analysis in families segregating the disease phenotype. This approach has been greatly enhanced by the generation of a high resolution genetic map of the human genome with very informative microsatellite polymorphisms by Genethon and other groups (Weissnebach et al., 1993). Using these genetic resources, a large number of disease genes have been placed on the chromosome map. The density of the genetic map and the success of positional cloning for monogenic diseases has spurred interest to isolate genes for more complex diseases such as diabetes, hypertension, obesity and psychiatric disorders. This is the major focus of the new Welcome Trust Centre for Human Genetics in Oxford in which the ICRF Human Genetics Laboratory is involved.

The research in the Human Genetics Laboratory involves two major projects:

A. Physical mapping of the human X chromosome

We have developed a number of resources in collaboration with The Genome Analysis Laboratory (Hans Lehrach, ICRF, London) to construct a physical map of the human X chromosome. One of the major goals is to isolate a series of overlapping human DNA fragments along the X chromosome. These overlapping contigs of clones have made it much easier to identify genes responsible for inherited disorders and cancers and are the first step in the human genome project objective to sequence whole chromosomes, To achieve this aim we are
using a system of isolating large fragments of human DNA (100-2000 kb) as linear artificial chromosomes in yeast cells. Yeast artificial chromosomes (YACs) allow rapid chromosome "walking" from a DNA marker genetically linked to a disease to the area containing the gene of interest.

B. Isolation of genes and identification of protein products for inherited disorders and cancers,

We are currently using the following methods for detecting new genes in cloned DNA. Once the region containing the gene of interest is isolated in a set of YAC clones, experiments are then done to identify and isolate candidate genes. In order to isolate new polymorphisms and expressed sequences, YAC clones are often represented by contigs of overlapping bacteriophage or cosmid clones in bacterial hosts from which DNA is easily purified. This is accomplished by subcloning the YAC after partial restriction enzyme digestion into these vectors and identifying human positive clones or by directly hybridizing YAC inserts purified from the yeast host chromosomes to gridded chromosome-specific cosmid libraries.

Various methods are then used to isolate expressed sequences including CpG island identification, searching for exons by conservation between species, exon trapping vectors, cDNA selection procedures and whole YAC and cosmid hybridization to cDNA libraries (for review see Monaco 1994). Once genes have been isolated, the nucleotide sequence is determined to predict the protein product of the gene and to identify mutations giving rise to disease in patients. For some genes, antibodies are produced that recognize the protein product of the gene thus allowing the study of its cell biological function. The normal and pathophysiological role of the protein product can then be related to the pathology and clinical symptoms seen in patients and more logical steps can be made towards therapies.

Some of the X-linked diseases that are being targeted for study in this laboratory in collaboration with other groups are in two distinct regions,


References

GENES RELATED TO HUMAN PRIMARY HEPATIC CANCER

Jian-ren Gu

National Laboratory for Oncogenes and Related Genes
Shanghai Cancer Institute
Ln. 2200 Xie-Tu Rd, Shanghai, 200032. PRC

I. Oncogenes activated in human PHC

Since 1983, at least seven proto-oncogenes and related genes have been demonstrated in this laboratory to be activated in human primary hepatic cancer (PHC), i.e. N-ras, c-myc, c-ets-2, IGF II, IGF II receptor (IGF II R), IGR I receptor (IGF I R) and c-fms or CSF 1 receptor (CSF 1 R). N-ras has been previously identified as a transforming gene from some PHC using DNA transfection assay [1]. Though the mutation of N-ras of PHC was not as frequent as expected, over-expression of its mRNA or protein product (p21) was frequently observed in about 80% of the PHC cases [1]. In addition, expression of c-myc and c-ets-s, putative transactivators for certain cellular genes, were also over-expressed in 80% and 100% of PHC samples respectively [2]. Furthermore, IGF-II and its receptors (IGF II R and IGF I R) and c-fms (CSF-1-R) genes were unexpectedly over-expressed both at mRNA and protein level in nearly all the PHC samples so far examined. As these growth factor and receptor and receptor genes were closely linked to cell growth, an autocrine/paracrine mechanism has been thus postulated to be involved in the autonomous growth of human liver cancer cells [3,4].

II. Anti-oncogenes for human PHC

1. p53 gene

In recent years, a survey of anti-oncogenes involved in human PHC has been undertaken in this laboratory. Among the well known anti-oncogenes, p53 gene has been reported to be involved in many types of human malignancies. In 1992, it was reported that the mutation at condon 249 of p53 gene has been observed in...
50% of PHC samples collected from Qi-Dong, China [5] and South Africa [6] An independent study from our laboratory demonstrated that mutation at codon 249 of p⁵³ gene was not necessarily essential of human hepatic cancer, but possibly related to some environmental factors [7,8]. Mutation at codon 249 of p⁵³ gene did occur in 5 out of 12 PHC samples from Qi-Dong, China, a high prevalent area for PHC, but it was observed in only 1 of 18 samples collected from Shanghai area. It implied that the mutation of codon 249 might not be unique for human PHC, but might be an event linked to some environmental factors, such as AFB1. However, loss of heterozygocity (LOH) could be frequently observed in PHC both from Qi-Dong (3 out of 12) and Shanghai area (6 out of 18), probed by pYNZ 22. Since pYNZ 22 should be a VNTR sequence, possibly located telemeric to p⁵³ locus, it was implicated that a putative novel anti-oncogene might be located in these loci.

2. Transthyretin (TTR) gene, a putative anti-oncogene in human PHC

Since 1989, subtractive cDNA libraries of normal human liver from PHC were constructed, cDNA clones have been analyzed by differential hybridization, Northern analysis and DNA sequencing. One of these clones has been characterized as transthyretin (TTR) gene [9]. In Northern analysis, the expression of mRNA was suppressed in most of the PHC samples. Furthermore, deletion of DNA was detected in 40-50% of PHC samples by Southern hybridization. Therefore, it implicated that TTR gene might be impaired in human PHC.

Recently, it has been demonstrated in our laboratory that the growth of human hepatoma cell line was retarded after transected with TTR gene. Moreover, the transected cell was sensitive to all trans-retinoic acid in further inhibition on the growth of hepatoma cells in vitro. These data implied that TTR gene might be a candidate of novel cancer-suppressed gene for human PHC [10].

III. Search for candidates of novel oncogene and cancer suppresser genes

1. Sequencing and characterization of clones from subtractive cDNA libraries
Subtractive cDNA libraries from normal human liver versus hepatic cancer and vise versa were constructed. More than 200 randomly selected clones have been sequenced. Eight clones with sequence not matched with GenBank were further characterized by Northern and Southern analysis. Deletion of genes have been identified in PHC DNA samples using 4 individual cDNA clones as probe. One of these clones, N_A_B (2kb) was sequenced. representing the 3' UTR of a cDNA (5kb) of a new gene highly expressed in brain tissues (Cheng Y., et al, to be published).

The sequence of a cDNA clone HP8 had been analyzed and it might be a new member of C/EBP family, with high mRNA expression in 2/5 PHC and 1/5 cancer-surrounding liver tissues (Xu L. X., et al, to be published).

2. mRNA differential display (mDNA DD)

Using mRNA DD techniques, near 30 cDNA clones have been cloned and sequenced. Except one (homologous to nucleoplasmin), all these clones have DNA sequence not matched with DATA BASE. The further characterization of these clones is in progress (Wan D. F., et al, to be published).

Reference

GENES INVOLVED IN ACUTE PROMYELOCYTIC
LEUKEMIA (APL)

Zhu CHEN. Sai-Juan CHEN
Shanghai Institute of Hematology, Key Laboratory of Human Genome Research of
National Ministry of Public Health. Shanghai Municipal Laboratory of Human
Genome Project, Rui-Jin Hospital. Shanghai Second Medical University.
Shanghai 20025, P. R. China

N-on random chromosomal translocations play an important role in the
pathogenesis of human hematological malignancies. These translocations result in
either deregulation of oncogenes or fusion genes with altered function. Acute
promyelocytic leukemia (APL) is one of the common leukemia subtypes
encountered in hematology, with vast majority of leukemic cells blocked at the
promyelocytic stage of granulocyte differentiation. APL is also characterized by a
specific, balanced chromosomal translocation t(15;17) (q23;q12-21) which is
present in most patients but never found in other types of leukemia, indicating the
importance of this translocation in leukemogenesis. Another character of APL is the
response of leukemia cells in vitro and in vivo to the differentiation inducer all-trans
retinoic acid (ATRA). ATRA is a naturally existing active metabolite of vitamin A
and has been shown to be able to induce the differentiation and inhibit the
proliferation of malignant cells in a variety of experimental systems. The successful
treatment of APL with ATRA has opened new prospects for the cancer
differentiation therapy.

The specific effect of ATRA in APL suggests that the retinoic acids and
their receptors should be important in the regulation of the differentiation,
proliferation and cell death of promyelocytes. Since 1987, important progress has
been achieved in the basic research on the retinoic acid receptors, especially the
discovery of two families of receptor: retinoic acid receptor (RAR) and retinoid X
receptor (RXR). Each family includes 3 members: α, β and γ. Sequence analysis
showed that RAR and RXR, like other members of the nuclear hormone receptor superfamily, bear six functional domains (A through E): A/B domain is responsible for ligand non-dependent transactivation. C is for the DNA binding, E is the ligand-binding domain with ligand-dependent transactivation property. These receptors are thus hormone-inducible transcription factors which regulate the target gene expression through binding to the retinoic acid response element (RARE) of these genes.

The chromosomal localization of RAR α (17q21) is very interesting since it is in the close vicinity of the generally accepted chromosome 17 breakpoint (1 7q12-21). Several laboratories, including ours, then took RAR α as one of the candidate genes involved in t(15;17). Indeed, when cDNA probes were used to perform the Northern analysis, aberrant mRNA bands were revealed in leukemia cells from some APL patients. Subsequently the use of genomic DNA confirmed the alteration of the RAR α gene structure. Interestingly, the rearrangements of RAR α gene have been mapped in almost all patients in the intron separating exons encoding the A and B domain of RAR α. By the simplest way of positional cloning, several laboratories have cloned the locus on chromosome 15 which is fused to the RAR α. This gene is now designated as PML (for Promyelocytic Leukemia). Our data show that PML gene spans about 50 kb and contains 10 exons. The 5' moiety of PML, corresponding to exons 1, 2 and 3, contains a proline-rich transactivation domain, a ring finger domain (atypical zinc-finger of C3HC4 type) involved in DNA binding and a coiled-coil domain responsible for protein-protein dimerization. The 3' moiety (corresponding to exons 4 to 6 and the alternatively spliced 3' end exons 7 through 10) of the gene is encoding an α-helical region. The gene has complex splicing pattern, mainly occurring in the region downstream of exon3. It may be interesting to note that PML belongs to a new family of nuclear protein including two other genes involved in malignant transformation, Raf and Ret. In APL, the chromosome 15 breaks have been mapped, in different patients, to two major regions with respect to the gene’s exon-intron structure, one is located in intron 3, exon 4 or the 5' end of intron 4, and the other is downstream of exon 6. The PML-RAR α fusion
gene transcripts are of different sizes according to PML rearrangements. The long (L)-type isoform is constituted by exons 1 to 6 of PML and the sequence encoding B through F regions of RAR α. The short (S)-type isoform includes only 3 PML exons (exon 1 to 3) and the same set of RAR α exons as in the L-type. In rare cases (about 5% in our series), the breakpoints have been mapped within exon 6 of PML.

The clinical relevance of the discovery of PML/RAR α is multiple. Using it as specific marker of the disease, we and others have shown that on day 10 to day 20 of the ATRA treatment for remission induction, there is an accumulation of matured granulocytes in the bone marrow which are still carrying the PML/RAR α rearrangement. This provided strong evidence that ATRA triggers APL cell differentiation in vivo. Perhaps the most important clinical application is the establishment of reverse transcriptase/polymersae chain reaction (RT-PCR) analysis of fusion transcript for the diagnosis and monitoring of the minimal residual disease during the clinical remission.

The discovery of PML-RAR α fusion gene has also provided important clue to the understanding of the pathogenesis of APL as well as the mechanisms underlying the responsiveness of APL cells to ATRA. In vitro studies established that PML-RAR α behaves differently than the wild-type RAR α in terms of RARE binding and transactivation. In some experiments it has been shown to be able to antagonize the function of normal RAR α. When introduced into HL-60 cells, PML-MR α block the differentiation induced by low concentration of ATRA. Like wild-type RAR α, PML-RAR α can also form the heterodimer with RXR. As RXR form active heterodimer with all nuclear hormone receptors including RAR, it is proposed that the sequestration of RXR by PML-RAR α could block the differentiation pathways normally regulated by other hormone receptors as well. In agreement with this, the expression of PML-RAR α in U937 cells abrogated the Vit D3 mediated differentiation. Recently it has also been shown that PML-RAR α can
disrupt the normal PML localization while ATRA can restore the normal PML subcellular distribution.

In the investigation of a large series of APL patients in China we have been able to discover a new chromosomal translocation t(11;17)(q23;q21). The cytology of this patient is a typical AML-M3. Clinically, this patient seemed to respond relatively poorly to ATRA since the bone marrow examined in day 19 of ATRA treatment revealed still a predominance of leukemia cells. Molecular cloning showed that RAR α was fused to a new transcriptional unit on chromosome 11q23. As this new gene contain 9 zinc tigers, it has been named PLZF (for promyelocytic Leukemia Zinc Finger). Analysis of PLZF genomic DNA revealed that this gene is at least 90kb long and contains 7 exons. Interestingly there is an usual splicing within exon 2, deleting a stretch of 369 bp and resulting the shorter A isoform of the PLZF transcript. In addition this splicing uses an untypical donor signal GA instead of GT whereas the acceptor signal is normal (AG). PLZF belongs to the Kruppel gene family. It has been shown to share an important homology at the N-terminal with the Bcl-6 gene which is rearranged in diffuse B cell lymphoma. Transection assay showed that PLZF-RAR α antagonize the transactivation of wild-type RAR α and RAR-RXR heterodimer. PLZF displays a much limited spectrum of expression compared with PML. Within the hematopoietic system, PLZF is mainly expressed in the early myeloid cells and is down-regulated during the ATRA induced myeloid differentiation. Preliminary work demonstrated PLZF-RAR α can also form heterodimer with RXR. It is thus possible that PLZF-RAR α block the promyelocytic differentiation in a similar way as PML-RAR α. This finding further confirmed the point of view that disturbance of the RAR-RXR regulated pathway plays a crucial role in leukemogenesis.

Recently, it has been shown that t(11;17) with PLZF-RAR α is a recurrent chromosomal accident in APL. An international collaboration of 6 laboratories examined 6 cases of APL with t(11;17). All 6 cases had the PLZF-RAR α fusion. The clinical data showed that these patients were resistant to ATM treatment.
when tested in vitro. cultured cells from three of them failed to differentiate in response to ATRA. It has been thus concluded that APL with t(11:17) and PLZF-RAR \(\alpha\) fusion gene is a discrete clinical syndrome with a worse prognosis than 1(15:17) APL. The comparative study between t(15:17) and t(11:17) APL will give new insights into the mechanisms of ATRA induced differentiation.

Our group is now working on the target genes of the RAR-RXR during the ATRA induced APL cell differentiation. By means of differential display, we have been able to obtain a number of cDNA clones from genes whose expression is induced or repressed after different times of treatment with ATRA. The study on the structure and function of these genes should be very important to understand the molecular and cellular basis of the differentiation therapy of APL with ATRA.

Acknowledgement: This work was partly supported by the National Natural Sciences Foundation of China, the Chinese High Technology Program (863) and Shanghai Foundation of Natural Sciences.

Main References
1. Chen Z, Chen SJ. Leukemia Lymphoma 1992; 8:253-260
MOLECULAR GENETIC BASIS OF Amyotrophic Lateral Sclerosis

Han-Xiang Deng, M. D., Ph. D.
State Key Laboratory of Medical Genetics
Hunan Medical University, Changsha, China
Department of Neurology, Northwestern University
Medical School, Chicago, U. S. A.

Amyotrophic lateral sclerosis (ALS), also called Lou Gehrig’s disease, Charcot’s disease, or motor neuron disease, is a paralytic disease due to the degeneration of large motor neurons of the cortex, brain stem and spinal cord. It causes progressive wasting and weakness of skeletal muscles and usually results in death within five years after the onset of clinical symptoms. Despite more than a century of research, no cause of, cure for, or means of preventing this disease has been found. About 5-10% of ALS are familial (FALS). FALS, clinically indistinguishable from sporadic form, inherited as an autosomal dominant trait with age dependent penetrance, provides an opportunity to use molecular genetic techniques to investigate the pathogenesis of this fatal disease.

Genetic linkage studies with more than 100 polymorphic markers from whole human genome on 23 kindreds established a locus for FALS on chromosome 21q22.1 and genetic-locus heterogeneity was also found (1). Subsequently mutations in Cu/Zn cytosolic superoxide dismutase (SOD1) gene were identified in FALS families linked to chromosome 21 locus (2,3). A total of 22 different mutations involving 17 distinct codons have now been identified in more than 100 families with FALS worldwide. These mutations occur in exon 1, 2, 4 and 5 of SOD1 gene, but not in exon 3 which form the active site loop of this enzyme. Mutation in exon 1 at codon 4 which changes the codon from GCC to GTC acid amino and from alanine to valine is the most frequent mutation that accounts for almost half of the mutations identified in FALS.
We mapped mutation sites onto the human SOD crystallographic structure to define the effect of these FALS mutations on the protein structure and function. The wild type human SOD structure was used to assess the contributions of the side chains at these positions to the active site, the dimer interface, and the β-barrel subunit fold. Overall, the structural results showed that the mutated side chains cluster near the two Greek key connections closing off the ends of the β-barrel, in the dimer interface, and at the base of the active site loops. These side chains, which are mostly sequence-conserved, are also structurally conserved in the wild type SOD structures from different species and appear to be critical for the structural integrity of the dimeric enzyme. Crystallographic structural studies predicted that the mutations in FALS produce less stable and therefore less active SOD enzymes which may play a crucial role in the pathogenesis of ALS. Enzymatic studies confirmed this prediction, the mutated enzymes have short half lives and less activity.

We made a transgenic mouse model with expression of a human FALS mutation at codon 93 of exon 4, which changes the codon from GGT to GCT, and amino acid from glycine to alanine. The transgenic mice developed paralytic, neurodegenerative disease resembling the human ALS phenotype (4).

Cu/Zn cytosolic superoxide dismutase is a well characterized “housekeeping” enzyme present in most aerobic organisms. It catalyses the conversion of the toxic superoxide anion radical \( \cdot \text{O}_2^- \) to hydrogen peroxide \( (\cdot \text{O}_2^- + \cdot \text{H}_2\text{O} + \cdot \text{O}_2\) \) which can then be converted to water by glutathione peroxidase or catalase.

The mechanism(s) through which the mutations in SOD gene cause ALS is largely unknown. However, the fact that these mutations destabilize the homodimer structure of SOD protein may be quite suggestive.

1) The destabilized proteins have short half lives and less activity, which will increase the oxygen free radical stress and facilitate the formation of strong
oxidant peroxynitrite (ONOO') form O$_2^-$ and notric oxide (NO). The neurotoxicity of peroxynitrite and its decomposition into toxic chemical species have been shown in vivo and in vitro. It has also been shown that the inhibition of SOD activity resulted in the apoptotic degeneration of neurons in vitro.

2) The mutations in SOD$_1$ gene confer the destabilized proteins new function(s) through some unknown mechanisms as the transgenic mouse model suggested. The elucidation of these mechanisms and the pathogenesis of FALS caused by the mutations in SOD1 will provide the ways for prevention and treatment of this disease.

Main references


Human Genetic Diversity and Its Evolutionary Implication
Takashi Gojobori and Tadashi Imanishi
DNA Research Center, National Institute of Genetics. Mishima 411, Japan

In elucidating structures and functions of the human genome, it is of particular interest to measure the genetic variability of human population at the molecular level. In particular, as the application of the human genome research to medical science is going to proceed further, it will be more important to understand the degrees and patterns of diversity in the human genome. Moreover, we may have to consider not only genetic diversity within a population of a particular ethnic group but also genetic diversity among populations of different ethnic groups, because it is said that there are more than 700 recorded ethnic groups in the world as of today.

To study the genetic diversity between different populations can also lead us to the understanding of evolutionary history of human populations through the genomic variation. Although it is obviously impossible to determine the nucleotide sequences of all individuals belonging to different ethnic groups, we may be able to measure the genetic diversity of human populations if we examine genetic markers suitable for this kind of study.

Genes in the human major histocompatibility complex (MHC) are typical examples of such suitable markers in the human genome diversity research, because some of the MHC genes exhibit an extremely high degree of polymorphism. The MHC molecules in humans are called human leukocyte antigen (HLAs), which were discovered in 1958 by J. Dausset. The HLA gene complex in the human genome is about 4Mbp long and is located on chromosome 6. The HLA genes, in general, are classified into two groups, classes I and II, depending on their structures and functions, although complement loci that are located within the HLA gene complex are sometimes called as class III. It is well known that the HLA-A, HLA-B and HLA-C genes in class I and the HLA-DR and HLA-DQ in class II are highly polymorphic for ethnic groups studied so far.
In 1991, the Eleventh International MHC Workshop was held at Yokohama, and the serological and DNA typing data were collected from 242 laboratories over the world. By serological typing methods a total of 17,959 and 5,770 individuals were examined for HLA genes and complement loci, respectively. By DNA typing methods, a total of 12,887 individuals were studied. If the overlaps were excluded, the data for a total of 27,765 individuals were collected in this workshop, and the number of ethnic groups examined was 130. All the data collected were stored in the MHC database that we have developed for the workshop.

By surveying the MHC database, we found that more than 15,000 individuals among a total of 27,765 came from unrelated and healthy individuals. We thus considered that the typing data for these individuals were suited for the analysis of genetic diversity. Moreover, the number of individuals examined for each of 84 ethnic groups was so large that we successfully estimated gene and haplotype frequencies for these ethnic groups.

We used the maximum likelihood (ML) method to estimate gene and haplotype frequencies, because the ML method is statistically more satisfactory than any other methods. For haplotype frequencies, we estimated the frequencies of not only two-locus haplotype but also three-, four-, and five-locus haplotypes. This enabled us to identify predominant MHC haplotypes for each ethnic group. We also calculated values of linkage disequilibrium for every haplotype. Statistical tests on linkage disequilibrium between any two loci were also carried out. Using these data, we studied geographical distributions of gene and haplotype frequencies in the world.

By use of the gene frequencies for HLA and complement loci, we prepared three different sets of gene frequency data, in order to calculate genetic distances. The first set includes data of two serologically typed class I loci, HLA-A and HLA-B, for 77 ethnic groups. We excluded the data of the HLA-C locus from our analysis, because the frequency of unknown alleles in the HLA-C locus was remarkably high for many ethnic groups. The second set includes the data of three
DNA-typed class II loci, DRB1, DQA1, and DQB1, for 20 ethnic groups. The third set includes data of three complement loci, Bf, C4A and C4B, for 37 ethnic groups.

Using these sets of allele frequencies, we calculated genetic distances, which measure genetic diversity between populations, for all possible pairs of ethnic groups. Phylogenetic trees for ethnic groups were constructed using the obtained genetic distances by the unweighed pair grouping (UPG) and neighbor-joining (NJ) methods. For the UPG method, it is assumed that the evolutionary rate is constant over time. On the other hand, this assumption is not required for the NJ method. However, the phylogenetic tree constructed by the NJ method gives only a network of ethnic populations without a 'root' of the tree.

The phylogenetic tree for allele frequencies of HLA-A and HLA-B constructed by the UPG method showed that 77 ethnic groups examined were classified into six distinct groups that roughly correspond to ‘Africans’, ‘Europeans’, ‘Oceanians’, ‘Native Americans’, ‘North Asians’, and ‘South-east Asians’. The ‘African’ group is separated first from the other four groups. The ‘Oceanian’ group includes such geographically distant populations as Australian Aborigines, Taiwan Aborigines, Yakults, and Inuits. The ‘North Asian’ group is composed of populations in northern China, Mongolians, Japanese, Koreans, and Tibetans, whereas the ‘South-east Asian’ group is composed of populations in southern China and South-east Asia. The phylogenetic tree constructed by the NJ method showed virtually the same features as that by the UPG method.

The phylogenetic tree constructed based on three DNA typed loci, DRB1, DQA1 and DQB1 showed a bit different feature of genetic relationships between various ethnic groups, compared with the phylogenetic trees based on the serological typing data. In the case of phylogenetic trees for the DNA typing data, only class II loci were used and the number of ethnic groups available was only 20. Thus, these factors may have influenced a difference between two kinds of trees.
The phylogenetic trees for three complement loci, Bf, C4A, and C4B, supported the genetic relationships among various ethnic groups shown by the phylogenetic trees based on the serological typing data, although the boundaries between major clusters of ethnic groups were less clear.

Finally, we studied the evolutionary process of the gene organization of the class I loci in the human genome. We collected 94 DNA sequences of class I MHC genes for human and other primate species from DNA databases, constructed molecular phylogenetic trees for them, and estimated the evolution of the class I multigene family. As a result, we found that the present gene organization of the class I loci in the human genome were created by repeated gene duplications and inversions, Such a dynamic evolution of the human genome must have taken place even after the divergence between Old World monkeys and New World monkeys.

From these studies, we concluded that the MHC genes were very useful for examining the genetic variability of human populations and the processes of genome evaluation.
FOURTEEN-YEARS STUDIES ON THE GENETIC DIVERSITY OF
ETHNIC GROUPS IN CHINA

Du Ruofu
Beijing 100101. China

China is a multinational country with an ancient civilization and a population of more than 1.2 billion. In China there are 56 identified (by Government) ethnic groups, as well as some 10 unidentified groups. Among these there are many communities which have been isolated for hundreds or even a thousand years, and some are unaware of their own origin. Moreover, China has a vast territory with a great variety of natural conditions and levels of economic development, as well as different cultural background in different areas. Thus, China provides a valuable natural laboratory for the study of human genetic diversity.

The laboratory of Human Population Genetics in the Chinese Academy of Sciences was established in 1980 and has been studying the genetic diversity of ethnic groups in China for 14 years.

The genetic diversity of 35 ethnic minorities have been studied. They are: Manchu, Korean, Hezhen, Mongolian, Daur, Ewenki, Oroqen, Huizu, Uygur, Kazak, Tibetan, Qiangzu, Yizu, Baizu, Hani, Lizu, Vazu, Naxi, Jingpo, Blang, Achang, Pumi, Nuzu, De'ang, Drung, Miaozu, Bouyei, Dongzu, Zhuangzu, Yaozu, Mulam, Maonan, Jingzu, Tujia, and Shezu. One unidentified group, "Benren" in Yunnan Province, and 11 subpopulations of Han from Hainan, North China, Inner Mongolia, Heilongjiang, Shanxi, Sichuan, Henan, Gansu, Fujian, Guizhou and Hakkas from Guangdong were also studied by us.

The genetic markers studied by us for exploring genetic diversity are as follows:
1. Red cell blood groups: ABO, MNSs, P, Rh, Lewis, Duffy,
Diego, Kidd, Ken, Xg, Letheran and ABH secretion

2. Serum proteins: Group-specific component, 3rd, 6th and 7th components of complement, haptoglobin, transferrin, and α₁-antitrypsin.

3. Red cell enzymes: Acid phosphatase, esterase D, phospho-glucokinase-1, glutamate-pyruvate-transaminase, 6-phosphogluconate-dehydrogenase, glyoxylase 1, adenosine deaminase, aldehyde dehydrogenase and lactase.

4. Other markers: PTC (phenylthiocarbamide) taste blindness, earwax, red-green color blindness, etc.

It is impossible and unnecessary to enumerate the gene frequency data of each genetic marker obtained by us from all populations listed above. Only some interesting results and conclusions are presented as follows.

1. Han has a very great genetic diversity

The Han, as the largest ethnic group in the world, makes up about one-fifth of all mankind. It is clearly shown by the genetic distances calculated and the dendrogram constructed on the basis of their gene frequency data, as well as surname frequency data and physical characteristics, the Han people can be obviously divided into northern and southern types.

The Hans living in the Yangzi River valley are of middle, or mixed type between the northern and southern Hans, but are more like the northern.

As we know, the Han originated from the Huaxia tribe-group which formed during the 21st-8th Century B.C. in the middle and lower reaches of Yellow River. When the Huaxia developed and spread to all directions, it absorbed many other ethnic groups, either partially or fully, through the latter’s announcement of changing into Han on their own accord, or sometimes by ruler’s order. In fact, the
major blood of today’s local Hans in many border areas of China, such as Hainan, Guangdong, Guangxi, Fujian, Northeast China, etc. came from the local ancient ethnic groups. but not Huaxia. One of the strong evidences for this is that the genetic distances between the Han and their neighboring local ethnic minorities are usually closer than those between Han subpopulations in different areas. for example, between Guangdong Hans and Northeast Hans.

2. Tibetans originated from the north.

According to the dendrogram and principal axes analysis made on the basis of genetic distances it can be conclusively said that the Tibetans are close to northern ethnic groups in China and are originated from the north.

3. Uygurs are Mongoloids.

The Uygurs in Xinjiang have fair skin, high nose bridge, deep set eyes, light eyes and hair, and can easily be mistaken for Caucasoids. Actually, the Uygurs belong to the Mongoloid race, but have some Caucasoid blood. They are genetically close to northern Mongoloid ethnic groups in China, but very far from Caucasoids.

4. "Benren" have mixed greatly with local ethnic groups in Yunnan Province.

In 1992, some Khitan small characters were found on the Benren’s tombstone. Studies on surnames, genealogy, and history followed then also confirmed that Benrens are descendants of Khitan, an ancient ethnic group in Northeast China, which established the Liao Kingdom (947-1125 A.D) there. After the Western Liao Kingdom (1131-1218 A.D) was subjugated by Mongols, Khitan disappeared in history. Our study on genetic diversity of "Benren" showed that they are genetically closer to the local ethnic groups Blang, Deang, Achang, Vazu than to Northeast ethnic groups in China.
5. Genetic diversity of adult lactase in Chinese ethnic groups.

The study on the lactose absorption among 6 subpopulations of Han and 7 national minorities showed that there were no significant differences in lactase phenotype distribution between the 6 subpopulations of Han, with an average of 92.4% of lactose malabsorbers. The Kazaks has a lowest frequency of lactose malabsorbers among national minorities, but still as high as 76.4%. Other ethnic groups show a gradient of lactose absorbers percentage from west to east and from north to south. It can be said that there was no strong selection in favor of the lactase persistence gene in China, and the west to east and north to south decrease in the frequency of lactase persistence is likely due to historical proliferation and spread of lactase persistence from the milk-dependent nomads in the desert zone of Southwest Asia.
1. China Rice Genome Project

As the world population is expending at an explosive rate, the demand for food is increasing dramatically. Rice is one of the main food crops in the world, especially in China. China has a long history of rice cultivation. Chinese scientists have studied it from various aspects for a long time. Based on the circumstances of China, rice has been chosen as the first organism for genome project,

Rice Genome Project was officially initiated in August of 1992. It is the biggest project supported by the State Science and Technology Commission of China for a single organism in the history of biological research in China. Professor Guo-fan Hong of Chinese Academy of Sciences and Professor Zhang-liang Chen of Peking University are in charge of the project. It invokes six laboratories nationwide with one center and five satellite laboratories,

The center is located in Shanghai, headed by Professor Guo-fan Hong. The five satellite labs are in the following institutions: Institute of Genetics, CAS; Central China Agriculture University and China National Rice Research Institute. The Center is mainly responsible for coordinating the project, constructing rice genomic libraries, setting up computer system, and etc. Two labs in Peking University and Fudan University are mainly involving in DNA sequencing and the three labs in other institutions are largely concentrated on RFLP mapping. In the first five-year plan, we intend to construct a YAC or BAC library, a RFLP map with a 3-5cm resolution; to isolate and sequence some organ-specific genes or genes with important functions; to set up chromosome-sorting system and computer analysis system. Through the effort of scientists, we have made some achievements in two years. Several labs will report their work separately in the conference.

II. Progress report of PKU lab
Reproductive period is very important in rice life cycle. Studying the rice flower-specific genes will not only have theory, but also application significance. We have isolated and sequences several flower-specific clones from cDNA libraries constructed from different floral organs at different developmental stages. Most of them do not have homologous sequences with the DNA sequences in the GenBank. In addition, we have cloned a calmodulin gene which involves regulating plant responses to light and a gene encoding protease which can inhibit the growth of certain pathogenic fungi in vitro. Both of them have been mapped onto a rice RFLP map, and the latter is located at a locus with the trait of resistance to rice fungal blast positioned by traditional mapping method. Two other genes important to quality of rice have also been cloned ---- ADP-glucose Pyrophosphorylase gene and 10KD prolamin gene. Now we are in the process of studying the function of genes mentioned above, especially the flower-specific genes.
RICE MOLECULAR MAPPING AND GENE LOCATION
Zhu, Lihuang, Shen, Lishuang, Lu, chaofu, Chen, Ying, Chen, Shouyi, Hong, Dejun, Xu, Yunbi*
Institute of Genetics, Academia Sinica, Beijing
(*Zhejiang Agricultural University, Hangzhou)

Ever since the first RFLP map of rice was constructed in 1988, great efforts have been devoted to developing permanent populations, including doubled haploids (DH) and recombinant inbred lines (RI) for further molecular mapping. Availability of such populations is also important for mapping quantitative trait loci (QTLs) and for international cooperative research. Using specifically designed culture medium suitable for Indica/Japonica hybrids, we obtained a DH population by anther culturing and subsequent chromosome doubling from a cross between Indica variety ZYQ8 and Japonica variety JX17. The DH lines developed from the cross show phenotypic stability through generations. For the most traits, phenotypic variance within each DH line is comparable to the mean phenotypic variance of each parent. The degree of transgressive segregation for most agronomic characters in the DH population is lower than that in the corresponding F2 population. Therefore this population can be permanently maintained for rice genome research.

A molecular linkage map based on this population has been developed and will be further developed towards saturation. At present, the map consists of more than 300 loci, including RFLPs, RAPDs and isozymes. Two previously unknown isozyme loci and a rice blast resistance gene, Pi-zh, were located on the map. By amplification of telomere associated sequences (TAS), three TASS were located at the extreme ends on this map, meaning that the physical ends of some chromosomes were reached in our genetic mapping. The graphical genotypes of all the DH lines were generated for haploid breeding programme by using HYPERGENE software.
Because ZYQ8 and JX17 are typical indica and japonica rice respectively, significant differences between them can be measured in ten agronomic traits and eight grain quality characters. Both of their F2 and DH population were used for mapping these QTLs. However, the use of DH population provides a way of evaluating identical genotypes in different environments, which make it possible to replicate phenotypic measurement. Four QTLs mapping approaches have been used for quantitative traits, including days to heading, plant height and grain weight.
TWENTY FIVE ETHNIC MINORITIES IN YUNNAN PROVINCE: THEIR GENOTYPIC AND PHENOTYPIC CHARACTERISTICS
Jiayou Chu
Institute of Medical Biology, Chinese Academy of Medical Sciences. Kunming 650107, P. R. China

There are 56 nationalities in China and 26 nationalities (including Han nationality) distributed in Yunnan Province, each having independent inhabitation areas. And some of them are genetically isolated populations. In the source of nationalities and genetic phenotypes, each nationality has its unique characteristics. There are significant differences in categories, incidence of genetic diseases.

1. General information of Yunnan Minority Nationalities

<table>
<thead>
<tr>
<th>Name of Nationalities</th>
<th>Population</th>
<th>Name of Nationalities</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yi</td>
<td>4,054,177</td>
<td>Zang(Tibetan)</td>
<td>111,414</td>
</tr>
<tr>
<td>Bai#</td>
<td>1,901,056</td>
<td>Jinpo#</td>
<td>118,322</td>
</tr>
<tr>
<td>Hani#</td>
<td>1,248,106</td>
<td>Bulang#</td>
<td>81,768</td>
</tr>
<tr>
<td>Dai#</td>
<td>1,014,318</td>
<td>Buyi</td>
<td>34,061</td>
</tr>
<tr>
<td>Zhuang</td>
<td>1,003,901</td>
<td>Pumi#</td>
<td>29,302</td>
</tr>
<tr>
<td>Miao</td>
<td>896,712</td>
<td>Achang#</td>
<td>27,613</td>
</tr>
<tr>
<td>Lisu#</td>
<td>557,144</td>
<td>Nu#</td>
<td>26,583</td>
</tr>
<tr>
<td>Hui</td>
<td>522,046</td>
<td>Jino#</td>
<td>17,843</td>
</tr>
<tr>
<td>Lahu#</td>
<td>408,203</td>
<td>De’ang#</td>
<td>13,399</td>
</tr>
<tr>
<td>Wa</td>
<td>347,738</td>
<td>Shui</td>
<td>7,688</td>
</tr>
<tr>
<td>Naxi#</td>
<td>265,708</td>
<td>Menggu</td>
<td>13,168</td>
</tr>
<tr>
<td>Yao</td>
<td>173,144</td>
<td>Man</td>
<td>7,055</td>
</tr>
<tr>
<td>Dulong#</td>
<td>5,536</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

# Those nationalities live in Yunnan province only.
II. Why So Many Minority Nationalities Are in Yunnan Province?

1. **Yunnan** Province is one of the places of human origin.

2. Special historical reasons in ancient China made many minority nationalities immigrate to Yunnan.

3. Complicated geographic environment, climate and rich natural resources in Yunnan Province fit for different minority nationalities’ living custom, so Yunnan became an ideal region for many nationalities to live in.

4. Yunnan Province was in a closed state to outside because of long-term feudal rule, regional chieftain domination and inaccessibility of transportation. These reasons kept many minority nationalities to be independent and isolated groups.

III. Minorities’ Physical characteristics

It’s observed that there exist some common physical characteristics and remarkable difference through somatometry investigation of adults (male, age from 24 to 60; female, age from 23 to 35) from 14 kinds of minority nationalities (Yi, Bai, Hani, Bulang, De’ang, Dai, Jingpo, Lisu, Miao, Naxi, Achang, Pumi and Lahu) in Yunnan Province.

1. Somatoscopy

   Common characteristics: Skin color is light yellow-brown; hairs are black and lissotrichous with less body hairs and beards. Direction of eyeslits for most people is in the level; eyefold of the upper eyelid grows well; minority nationalities living in South-West region of Yunnan Province (Hani, Jino, Dai, Jingpo, et al.) show a higher mongoloid fold rate, while those living in North-West region (Naxi, Pumi, Bai, Lisu, et al.) show a lower mongoloid fold rate. Most people have a
upper nasal tip and nasal base, especially for the females: Alae nasi develop well and breadth of the oral fissure looks wider.

2. Somatometry

Through analysis on male adults’ cephalic index (head length x100/head width), morphological facial index (face height x100/face width) and nasal index (nose height x100/nose width), we obtain the basic shape of head and face for minority nationalities in Yunnan Province: the length and width of head shares suitable proportion in mesocephaly pattern, face looks high and narrow in leptoprosopy pattern, nose looks share and narrow in leptorrhing pattern. Generally speaking, the basic characteristics are: appearance with narrower head breadth, face breadth and bigonial breadth matching with a wider breadth of the end fissure.

From the classification of body height, minority nationalities in Yunnan Province belong to sub-middle short pattern, among them, Zang and Pumi are the highest in height, Miao is the shortest in height; Naxi is the most heavy in weight. Miao is the most light in weight.

IV. Dermatoglyphic Study of Minority Nationalities

Using comparative method to study inheritable characters of dermatoglyphic in 11 nationalities (Jingpo, Yi, Dai, Lisu, Bai, Naxi, Dulong, De'ang, Achang, Nu, Lahu, Hani, Buiang, Jino and Pumi nationalities, by Anlu Jin). The result shows that the characters of dermatoglyphic not only exist significant differences between White and Black people, but also in different nationalities.

V. Minorities’ enzymology differences

There are some differences in RBC enzymes and HLA antigens among Yunnan minority nationalities, here is an example:
The phenotype distributions of esterase D (EsD), phospho-glucosomutase-1 (PGM1), glyoxalase I (GLOI), erythrocytic acid phosphatase (EAP), adenosine deaminase (ADA) and adenylylate kinase (AK1) were detected in 22 ethnic groups which are Yi, Bai, Hani, Zhuang, Dai, Miao, Lisu, Hui, Lahu, Wa, Naxi, Yao, Zang, Jingpo, Bulang, Pumi, Nu, Achang, De'ang, Jino, Buvi, and Dulong. The results showed that the range of genes frequencies of six RBC isoenzymes (including PGM1 subtype) in the 22 ethnic groups were as follows respectively.

- EsD 0.5223-0.7282
- PGM1 0.6196-0.8100
- GLOI 0.0833-0.2974
- EAP 0.0922-0.4474
- ADA 0.9107-0.9964
- AK1 0.9554-0.9955

(in 6 groups, in the other groups AKA was the same as 1.0000)

In addition, rare phosphoglucomutase phenotypes were found in 14 groups, the gene frequencies ranged from 0.0033 to 0.0350.

VI. Minorities’ genetic diseases

There are also great differences in genetic diseases among Yunnan minority nationalities, for example, the incidence of β-thalassemia is very high in Yunnan (5.51%), the Dai and Jingpo are two nationalities which have the highest incidence. Here are another two examples:

1. G6PD (glucose-6-phosphate dehydrogenase) deficiency gene rate:
   - De’ang 0.203
   - Hani 0.0370
   - Dai 0.161
   - Lahu 0.1136
   - Jingpo 0.045
   - Wa 0.0821

2. Rate of abnormal hemoglobin
   - Achang 42.925%
   - Aini 1.18%
   - Jino 16.25%
   - Bai 0.479%
   - De’ang 13.99%
   - Lahu 0.299%
   - Jingpo 12.25%
   - Wa 0.182%
VII. Mitochondrial DNA polymorphism and Random amplified polymorphic DNA

Using sixteen restriction endonucleases to assay the mtDNA in Buyi, Miao, Shui and Han nationalities (by Huaqian Jin). the result shows that Shui nationality possesses the highest polymorphism: the Han and Miao are the most closely related; and the genetic relation between the Buyi and the Shui is somewhat distant.

Using twenty-two 10bp primers of arbitrary nucleotide sequence to analyze random amplified polymorphic DNA (RAPD) in Lisu, Naxi, Bai and Zang minority nationalities (by Wen Wang), the result shows that Lisu nationality has higher genetic variation, whose mean index of genetic distance (1-F)is 5.465%, the Naxi has lower variation. average 1-F is 2.283%; Bai and Naxi are closely related (1-F is 4.421%).

VIII. Reservation of different Chinese nationalities’ genome in Yunnan and comparative studies of gene loci

At present, more and more youths are married to each other among different nationalities, genome of some nationalities face the danger of extinction, therefore, reservation of nationalities genome become an urgent project. If measures have not been taken from now on, we will lose the opportunity forever.

Our research includes:

1. Take the blood samples of different nationalities such as Dai, Jingpo, De’ang and Yi, prepare DNA.
2. Use the methods of PCR, DNA hybridization, and DNA sequencing etc. to make the gene diagnosis of thalassemia, G6PD deficiency, abnormal hemoglobin disease.

3. Comparative studies on different gene loci among different nationalities

4. Use the EB virus transformation technique to establish the genetic-specific immortalized lymphocyte lines of different nationalities in order to do further research in the future.

Acknowledgements: We would like to thank the following colleagues for offering data. Drs. Li Ming, Wang Wen, Jin Anlu, Peng Lin, Zou Langping, and Shen Bin.
ANALYSIS OF SOME RICE GENE SEQUENCE AND ANALYSIS OF RICE CHROMOSOMES

Cao Kai-ming Zhan shu-xuan. Sun Chong-rong
(Dept. of Biochemistry, Fudan University. Shanghai, 200433)

As a satellite laboratory of Chinese Rice Genome Research Program, our research area is concentrated on the analysis of rice gene fragments and analysis of rice chromosome which includes the construction of cDNA library and genomic library, sequencing of gene fragments, identification of each chromosome and localization of some gene on chromosome. Since the Genome Research Program was started in China, we have analyzed about 250 gene fragments from rice genome completely or partially, and all 12 chromosomes of rice have been identified preliminarily by using computer image analysis system, Some of these results will be presented as follows:

1. Construction of cDNA library and sequencing of cDNA clones.

Isolation and analysis of cDNA clones are indispensable for characterization of the expression mode of rice genome on specific growth stages or organs, and cDNA isolation will be very useful for making RFLP map or a physical map. "GuangLu Ai 4" rice species was used as basic research material that is the appointed cultivar for the “Chinese Rice Genome Research Program”. Rice mRNAs were prepared and purified from embryo, shoot and root respectively, in case of shoot and root, ten-day-old etiolated seedling was used. Pure mRNA were ligated to vector $\lambda$ zap II and three relevant cDNA libraries were constructed. So far, about 230 cDNA clones randomly selected from rice shoot cDNA libraries have been sequenced. After the nucleotide sequence data were translated to amino acid sequences, the results of similarity search indicated that about 100 kinds of protein homologous were putatively identified. Some examples of putatively identified
proteins are listed in Table 1. Now, we are going to accumulate cDNA sequence data further, and we are using these fragments for genome mapping.

2. Construction of rice genomic library and screening and analysis of some genes.

For physical mapping of rice genome and gene sequencing, a rice genomic library was constructed using EMBL3 as vector with original efficiency of $5.4 \times 10^5$ pfu. Total rice DNA was prepared from ten-day-old eliotated shoot by CTAB method, and the phage library was generated by ligation of partially Sau 3A-digested, size-fractionated and dephosphorylated DNA with the arms of the $\lambda$ EMBL-3 vector, This library will be used to support the other library (YAC, BAK etc.) for making a contiguous genome map. At this stage, applying cDNA fragment or those probes from other plants. Some clones including interested genes have been screened from this phage library by in situ plague hybridization. About 20 gene fragments have been isolated and sequenced including $\beta$-tubulin, U2 snRNA carboxypeptidase Y, actin, amylase, chitinase. From physical mapping and partial sequence of some clones, it can be seen that some gene families in rice appear varied organization pattern from other plants. For example, we have found 8 U2 snRNA genes in rice genome that is similar to U2 snRNA gene family in Arabidopsis thaliana but the rice U2 snRNA genes appear to be dispersed and are flanked by distinct upstream and downstream sequences.

3. The computer image analysis of rice pachytene chromosome

Rice diploid cell contain 12 pairs of chromosomes. There have been some considerable discrepancies in the identification, characterization and numbering of rice chromosomes, mainly due to their small sizes and similarity in morphological characteristics, the subjectivity of researchers, and perhaps, partially due to the inherent differences among cultivars.
Table 1: Putatively Identified Protein Sequences From Rice Shoot

<table>
<thead>
<tr>
<th>cDNA</th>
<th>Protein</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDRSC1</td>
<td>αβ-Gliadin storage protein</td>
<td>Wheat</td>
</tr>
<tr>
<td>FDRSC3</td>
<td>Ribosomal protein L41</td>
<td>Rice</td>
</tr>
<tr>
<td>FDRSC4</td>
<td>Ribosomal protein L41</td>
<td>Rice</td>
</tr>
<tr>
<td>FDRSC5</td>
<td>Phenylalanine ammonia-lyase</td>
<td>Rice</td>
</tr>
<tr>
<td>FDRSC6</td>
<td>Seed ricilin</td>
<td>T.cacao</td>
</tr>
<tr>
<td>FDRSC7</td>
<td>Tuber protein patatin B2</td>
<td>Potato</td>
</tr>
<tr>
<td>FDRSC8</td>
<td>Cytochrome b5</td>
<td>Rice</td>
</tr>
<tr>
<td>FDRSC9</td>
<td>Flavanone-3-hydroxylase</td>
<td>H.vulgare</td>
</tr>
<tr>
<td>FDRSC10</td>
<td>26s rRNA</td>
<td>M.sativa</td>
</tr>
<tr>
<td>FDRSC11</td>
<td>Glycerinaldehyde-3-phosphatase</td>
<td>Spinach</td>
</tr>
<tr>
<td>FDRSC13</td>
<td>α-Amylase</td>
<td>O.sativa</td>
</tr>
<tr>
<td>FDRSC14</td>
<td>Dihydrodipicolinate Synthase</td>
<td>T.aestivum</td>
</tr>
<tr>
<td>FDRSC15</td>
<td>Aspartic proteinase</td>
<td>Barley</td>
</tr>
<tr>
<td>FDRSC16</td>
<td>Phenylalanine ammonia-lyase</td>
<td>G.max</td>
</tr>
<tr>
<td>FDRSC17</td>
<td>Light-harvesting chlorophyll a/b-binding protein</td>
<td>Rice</td>
</tr>
<tr>
<td>FDRSC18</td>
<td>UFGT</td>
<td>Maize</td>
</tr>
<tr>
<td>FDRSC24</td>
<td>Oleosin BN-III</td>
<td>B.napus</td>
</tr>
<tr>
<td>FDRSC25</td>
<td>Sucrose-UDP glucosyltransferase</td>
<td>O.sativa</td>
</tr>
<tr>
<td>FDRSC26</td>
<td>Storage protein (UBC1)</td>
<td>Wheat</td>
</tr>
</tbody>
</table>

For the first time, we introduced computer image analysis system into the study of chromosomes of rice microsporecytes at pachytene stage. Owing to the lengths of chromosome at this stage is ten-fold longer than in somatic cell, more detailed morphological characteristics of rice chromosomes will be able to identify. We have obtained objective image data (named condensation pattern) which can
reflect the number and pattern of chromomere-band and other morphological characteristics of each rice chromosome more objectively, directly and accurately than the “string of beads” ideogram obtained by traditional methods.
SEQUENCING. IDENTIFICATION AND MAPPING OF
Schistosoma mansoni GENES
Sergio D. J. Pena (1). Gloria R. Franco (1). Manami Tanaka (2) and Andrew J.
G. Simpson (3)
1) Departamento de Bioquimica e Imunologia, Instituto de Ciencias Biologicas
-UFMG, Avenida Antonio Carlos,6627, 31270-901 Belo Horizonte.
MG, Brasil.
2) Institute of Basic Medical Sciences, University of Tsukuba.
Tsukuba, Ibaraki 305, Japan.
3) Laboratorio de Biologia Molecular, Centro de Pesquisas “Rene Rachou”-
FIOCRUZ, 30190-002 Belo Horizonte, MG, Brasil.
Correspondence to: Dr. S. D. J. Pena, Departamento de Bioquimica e
Imunologia, Instituto de Ciencias Biologicas-UFMG, Caixa Postal 486,
Belo Horizonte, MG 30161-970 Brazil.
Tel. (55-3 1) 227-3496; Fax (55-31) 227-3792;
e-mail: spena @vml. Icc, ufmg. br

Introduction

Although a decade has passed since the first Schistosoma mansoni
genes were cloned, relatively little information is currently available concerning
the parasite’s genome and its gene complement. Indeed, until 1993 GenBank
contained less than 100 different full-length cDNA or genome sequences from
schistosomes (GenBank 1993). In reality this means that our understanding of
host-parasite relationships as well as the search for potential targets of
chemotherapy and immunoprophylaxis are based on less than 1/0 of the total
potential genetic information. In the context of the global effort to
systematically map and sequence the human genome (Olson 1993), a program
that embraces the sequencing of a number of model organisms such as yeast,
Drosophila and Caenorhabditis elegans, we have thus embarked upon the
beginnings of a program to define the schistosome genome and its component
genes in much greater depth. We reasoned that application of genome methodology to gene discovery in *Schistosome mansoni* would considerably accelerate the slow pace of progress in understanding the molecular biology of this parasite.

A pilot program was started in 1992 as a product of collaboration between Sergio D. J. Pena and Andrew J. G. Simpson in Brazil and J. Craig Venter in the United States. An automatic fluorescent sequencer was installed in Pena’s laboratory and Expressed Sequence Tags (ESTs) were sequenced from a cDNA library constructed from mRNA extracted from adult worms. ESTs are single pass sequences of at least 150 bp derived from one or both ends of cDNAs which we obtain using automated DNA sequencers. ESTs provide an “identity tag of the gene” that can be used for the identification of the respective gene based on its homology with database sequences derived from the same or other organisms in the DNA data bases (Adams et al. 1991). The cDNAs from which the ESTs are derived then become important tools for selecting and identifying full length copies of the gene within DNA libraries and for mapping the position of the gene within the organisms genome.

An adult worm size-selected cDNA library was constructed and cloned into BA vector (a phagemid derived form pEMBL), essentially as described (Adams et al. 1993). The ESTs were obtained by partial sequencing from one or both ends of randomly selected cDNAs using M13 universal, M13-40 or M13 reverse primers, in automated DNA sequencers A. L. F. Pharmacia or 373A Applied Bio-system). Homology searches were done at National Center for Biotechnology Information site using the program Basic Local Alignment Search Tool (Altschul et al. 1990) for nucleotide sequences (BLASTN) or amino acid sequences (BLASTX). The genes were identified based on the high score alignments between the EST and the database sequences. The cDNAs were classified as “identified genes” when they showed matches with schistosome sequences or with those of other organism in the databases, or
“unknown genes” when they showed only partial homology with nonschistosome genes or had no significant database match.

We obtained a total of 607 ESTs from 429 cDNA clones. From these clones, 19.8% corresponded to undesirable sequences such as rRNA, mitochondrial sequences and plasmids without inserts. The remaining were grouped as follows: (1) 33.1% with no database match which may correspond to untranslated regions of the cDNAs or genes expressed only in *S. mansoni*; (2) 8.4% with a partial match with non-schistosome sequences and whose alignments consisted either of short regions probably corresponding to either conserved motif or domains or regions enriched with a specific amino acid; (3) 16.1% homologous to previously sequenced schistosome genes and (4) 22.6% showing homology with genes from other organisms. The first two groups consisted of “unknown genes” and the two last groups of “identified genes”. The ESTs with no database match were compared one with the other and grouped into high-homology clusters. Presumably these correspond to distinct genes, some of which are represented with high frequency in the library. The ESTs collected to date represent a broad range of transcripts including enzymes and known antigens as well as cytoplasmic, structural, regulatory, membrane, nuclear and secretory proteins.

The results were thus excellent: in approximately one year the number of known genes in *S. mansoni* more than tripled. Many interesting genes were identified, encompassing several transcription factors that have been already completely sequenced and whose biology is being intensely investigated. It is difficult to imagine how such genes would have been discovered in schistosomes were it not using the EST approach. A manuscript with the results of this pilot program is in press (Franco et al., 1994).

We have also initiated efforts to map the identified genes. *S. mansoni* has seven pairs of autosomal chromosomes and one pair of sex chromosomes (ZZ for a male worm and ZW for a female) and a haploid genome size of
2.7x10^8 basepairs (bp). In Tanaka’s laboratory, a yeast artifical chromosome (YAC) library was constructed with partial digested parasite genomic DNA and the chromosome location of each insert was detected by fluorescent in situ hybridization (FISH). The library contains 1283 clones with an average insert size of 358 kb, which represents 1.7 fold coverage of the genome (>4.6x10^8 bp). One hundred randomly selected YAC clones were localized by FISH and found to be distributed widely among all the chromosomes. The assembly of 14 YACs covered approximately the entire region of chromosome 3. The YAC library was then used for mapping of the genes identified by the EST strategy as described above. The mapping of a gene homologous to a human Y-box binding protein gene to the parasite’s chromosome Z represented the first instance of gene mapping in S. mansoni. A manuscript with these results is also in press (Tanaka et al., 1994).

In summary, our initial results with the genomic approach to the study of Schistosoma mansoni are quite promising. If successful, these new strategies will not only considerably further our molecular knowledge of this parasite with all its attending practical implications, but will also serve as tactics of choice in the investigation of the many other infectious and parasitic diseases that afflict mankind.

ACKNOWLEDGMENTS

We thank Dr. J. Craig Venter and the TIGR staff for their inestimable collaboration, Dr. M. Bento Soares from Columbia University in whose laboratory was constructed the cDNA library and D. H. Hirai from Kyoto University for the FISH work. This research was supported by TIGR, FAPEMIG and Fundacao Oswaido Cruz.

REFERENCES


GenBank 1993. Release 80, December 10


AUTOMATED DNA SEQUENCER WITH A NEW CONTINUOUS
DETECTION SYSTEM FOR SIMULTANEOUS ON-LINE
SEQUENCING OF 44 CLONES

European Molecular Biology Laboratory, Meyerhofstr. 1, 69117 Heidelberg.
Germany

A new high density continuous detection system was developed and used in
routine DNA sequencing, using both labeled primers and internal labeling
techniques with T7 DNA polymerase, as well as cycle sequencing protocols. The
present version allows simultaneous sequencing of 44 clones, each yielding around
1000 bases.

The sequencer retains the advantages of a non-scanning system, i.e. fixed
laser, no moving parts, simplicity, robustness, and high sensitivity. In addition, the
performance of the new system does not depend upon a straight running band
migration pattern, since there is a continuous row of detectors and two of them
cover each track. It permits the use of glass plates of lower planarity. This feature is
of importance for our ultra thin gel system ( thickness less than 200 µm ), which
allows the use of high voltages resulting in a high sequencing speed ( around 800
bases per hour per clone). It is also of particular importance with long gel
separation distances resulting in long read-outs ( around 1000 bases), improved
single peak resolution, high accuracy, sequence reliability, low redundancy and
efficient walking primer strategy.

The system allows the simultaneous on-line sequencing of two
complementary strands of a double stranded template in a single sequencing
reaction with two different primers, each labeled with a different fluorescent dye.
The two reaction products are loaded together, excited by two laser beams and
detected at two different positions in the gel. Since the sequences of both strands of
the template are determined simultaneously, the costs of labour, DNA template
preparation and sequencing reactions, as well as of gel casting, are reduced by a factor of two.

The miniaturization of the detectors and the reduction of the sample track size were made possible by improvements in the optical system, increasing the horizontal resolution while maintaining the high sensitivity.

Software is being developed for efficient handling of the high data throughput resulting from the large number of samples.
Towards a More Efficient Large Scale DNA Sequencing Strategy.
Harmut Voss, Jurgen Zimmerman, Josef Stegemann, Stefan Wiemann, Christian
Schwager and Wilhelm Ansorge.
European Molecular biology Laboratory. Meyerhofstr. 1, 69117 Heidelberg.
Germany.

An efficient DNA sequencing strategy should allow the determination of the
correct consensus sequence on both strands of an unknown DNA fragment using a
minimum number of sequencing reactions on fragments with minimum overlap.

For cosmid scale DNA sequencing we use at present a directed strategy
based on inexpensive primer walking, which has been applied to several cosmids
from chromosomes IX, XI and XV in the course of the European Community yeast
genome sequencing project. Double stranded DNA sequencing reaction protocols
using T7 DNA polymerase and internal labelling flourescein-15-*dATP allow
accurate low redundancy, automated, large-scale DNA sequencing with unlabeled
primers.

Inexpensive walking primers ( $8.00/primer are synthesized in sets of 10
(25 in the future ) on the EMBL multiple segmental DNA synthesizer. On poorly
growing cosmids cycle sequencing is performed using SequiTherm DNA
polymerase. The number of fragments required to determine the sequence of a
cosmid insert was further reduced by increasing the reading length per reaction of
up to 1000 bases with a single peak resolution of up to 700 bases in unprocessed
raw data on Hydrolink gels between 5% and 6% 60cm long gels.

The statistics of sequencing cosmid inserts by the directed primer walking
strategy were compared to those projects performed by using either random or
semi-directed approaches: while usually 800-1000 sequencing reactions are required
to finish a cosmid in a random shotgun project, less than 200 reactions are required
in the described directed approach. However, a number of about 150 primers have
to be synthesized for the directed primer walking strategy, making this approach more affordable when low price primers are available. Several ways to reduce the number of required primers are pursued:

(i) increasing further the reading length per primer:
(ii) bi-directional nested deletions on large cosmid inserts and;
(iii) increasing the number of initial standard primer runs during a limited shotgun phase.

On the new EMBL sequencing device we evaluate efficiency of a sequencing strategy, using two differently labeled “universal” and “reverse” primers in simultaneous sequencing of both strands on plasmid subclones.
The Embl Multiple (10-25 FOLD) Segmental DNA Synthesizer
Brian Sproat, Siegfried Winkler, Heiko Drzonek, Thomas Rupp, Hans Witmann,
Hans Floser, Leo Burger, Carol Stettner, Christian Schwager, Wilhelm Ansorge,
European Molecular Biology Laboratory, Meyerhofstr. 1, D-69117 Heidelberg,
FRG.

For medium or large scale DNA sequencing projects, ordered strategies,
offer the advantage of lower sequence redundancies. Among the ordered strategies,
the walking primer approach requires inexpensive, and reliable production of
large number of sequencing primers.

A multiple DNA synthesizer has been developed at EMBL, meeting these
requirements.

A user friendly control software has been developed. The structured
software allows the convenient modification of the synthesis cycle conditions
performed on the instrument. Various test routines enable easy maintenance of the
synthesizer. The user interface allows to enter the oligo sequences either manually
or to import oligo sequence files in various DNA sequence file formats.

Up to 10 oligos (25 in the near future), which may differ in length, can be
synthesized simultaneously. Additionally to the four bases, two further positions
are free for incorporation of other modified nucleotides. The cycle time is
approximately 6.5 minutes, i. e. ten 18-mers are synthesized within 1.5 hours. The
throughput of the synthesizer is about 60 oligos per day.

The synthesis scale can be reduced to 0.01 micromole scale, sufficient for
DNA sequencing. A reduced scale and the optimized valve/tubing layout result in
reduced consumption of chemicals (5 times lower compared to conventional
synthesizers) and in a reduced cost per base-coupling ($0.5/base, e. g. the synthesis
cost a 16-mer is $8.00).
A modified deprotection chemistry (ethanol/ethanolamine instead of standard ammonia chemistry) shortens the post synthesis steps to about 1.5 hours.

The oligos are produced with an overall yield of 98% per oligo (24-mer) on the segmental synthesizer. The oligos can be used for PCR after purification.

Since 1993 the multiple segmental DNA synthesizer has been routinely used in the EMBL DNA sequencing service as well as in the S.cerevisiae genome sequencing project.

The instrument is being modified for synthesis of up to 25 primers, thereby increasing the throughput to 150 oligos with one instrument in eight hours. Modifications in the layout and cycle chemistry will further reduce the price per base coupling.
Geneskipper: An Integrated Software Environment for DNA Sequence Assembly and Analysis in Mega Base Scale Sequencing Projects.

Christian Schwager, Stefan Wieman, Christoph Sensen, Thomas Rupp and Wilhelm Ansorge.
European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, FRG.

In the past years, automated fluorescent DNA sequencing has become a widely used tool in molecular biology and clinical diagnostics. Geneskipper integrates trace data from automated DNA sequencing, DNA sequence fragment assembly and DNA sequence analysis. Huge gene loci or even complete smaller genomes can be analyzed with Geneskipper.

Geneskipper's fragment assembly utility is able to handle up to 20,000 DNA fragments, each 64,000 bases in size, spanning a total consensus sequence of up to 64 mega bases.

To facilitate the project management and data handling in mega base projects, Geneskipper has the ability to organize the sequence information in "groups": DNA fragments from different clones or subclones may be arranged to form separate groups. The groups may be handled separately - like isolated subprojects - or different groups may be joined to analyze certain sub-regions or even the complete project at once. This allows fast and transparent analysis in large scale projects.

The sequence analysis utilities of Geneskipper - once reading frame search, pattern analysis, dotplot analysis, base composition analysis, oligo design - also work on sequences ranging up to 64 mega bases.

Geneskipper's contig editor can assemble and edit multiple alignments used for phylogenetic analysis, genotyping (e.g. HLA typing, viral strains). Up to 2500
individual sequences may be aligned. Distance and parsimony masks can be calculated automatically and written in standard multiple alignment formats (e.g. Phylip interleaved).

*Geneskipper* operates under Microsoft Windows 3.1 on IBM PC compatible computers. The program takes advantage of MS Windows support for any commercially available graphics adapters and printers.
Development of an Automated System for Loading of DNA Sequencing Gels.

Holger Erfle, Siegfried Winkler, Hans Wittmann, Josef Stegemann, Christian Schwager, Wilhelm Ansorge, European Molecular Biology Laboratory, Meyerhofstr. 1, D-69117 Heidelberg, FRG.

In recent years, most key steps in the DNA sequencing process have been automated with the aim to produce reliable and reproducible results. However, in most labs the loading of DNA samples to the sequencing gels is still done manually.

We have developed an automated system for loading sequencing gels. The high positioning accuracy of the motion system (0.01mm) allows to load standard DNA sequencing gels (0.35mm thickness) as well as ultra thin gels (down to 0.05mm). The loading volume ranges from 50 microlitre down to 0.5 microlitre. The constant even liquid flow ensures thin, highly precise starting zones of the DNA samples. Thirty two tracks are loaded in less than 8 minutes. The sequencing results obtained with the automated loader are better than those obtained with manual loading.

The control software allows adaptation to different geometries. Standard microtiter plates (96 well, or 60 well Terasaki plates may be used. Different gel sizes and comb geometries standard 5mm combs down to 1mm combs, or shark tooth combs) may be used.

One loader is used as a stand-alone system, the sequencing gel is loaded, pre-electrophoresed and then transferred to the automated sequencer. Another loading system is directly integrated into the sequencer. The transfer of the loaded gel is not necessary.
Besides the application to DNA sequencing the loader is also used in fragment analysis on polyacrylamide gels, as well as for loading of agarose gels (with a modified gel holder).

The automated gel loading system has been routinely used for two years in the EMBL DNA Sequencing service.
OVERVIEW OF GENOME SEQUENCING STRATEGIES

Ellson Y. Chen
Advanced Center for Genetic Technology, Applied Biosystems Division of Perkin-Elmer Corp., Foster City, CA 94404 USA.

The Human Genome Project in USA is a coordinated effort to discover all of the approximately 100,000 human genes and render them accessible for further biological study. Existing genetic maps are being merged with nascent physical maps of entire chromosomes, and ultimately the complete DNA sequence of the human genome will be determined. Sequencing technology thus becomes increasingly important.

Since the 1977 introduction of the two currently used sequencing techniques—dideoxy chain termination and chemical degradation—DNA sequencing has undergone rapid improvement. Today the dideoxy method has become the main procedure used in large-scale projects, and improvements over the past decade have included the introduction of various automated instruments employing high-power computers, signal detection devices and robotic workstations. Consequently a skilled investigator can determine the sequence of more than 100 kb per year, and a major production lab can thus produce over 1 megabase of sequence per year with >99.9% accuracy. Nevertheless, in order to elucidate the entire 3 billion bp of human genome within the next 10 years, a 10- to 100-fold improvement in sequencing throughput is still needed. Here we review current strategies and discuss the prospects for productive new methods.

Currently, since the length of sequence data obtained from a single experiment is limited to about 500 - 800 bases, determination of larger sequences requires a strategy to assemble the shorter ones. Two strategies, ordered and random, are used thus far. In the ordered approach, nested sets of deleted templates are ordered prior to sequencing, thus allowing the sequence data connected in sequential order. Advantages of this approach are (a) reduction of the
sequencing load by several folds: (b) better organization of raw data from individual
gel runs, and (c’) ability to sequence highly repetitive regions. However, the ordered
approach is tedious, difficult to automate, and thus far has only worked for inserts
of up to 5 kb in M13 or phagemid vectors. Auxiliary methods like primer-walking
(in which a new primer is made from the end of the sequencing tract) thus become
increasingly prevalent for this type of approach: but the primer-walking process is
inefficient -- making primers is slow and expensive, and primers may fail to anneal
with high-complexity DNA template. In one approach to solve this problem,
reusable tandem hexamers were employed as walking primers to decrease the cost,
but its consistency and suitability for automation remain to be proven.

As a result of these difficulties, the random approach is employed by most
large-scale sequencing projects, especially when carried out using high-throughput
automated sequences. Each sequencer is currently capable of analyzing 36 samples
of 550 bp per run (and more samples and longer readings per run are at hand).
Thus 40 kb or more of raw sequence data can be produced in a typical two-run-per-
day operation. A model has been provided by the nematode sequencing project,
producing over 6 megabases of contiguous sequence by random shotgun
sequencing of cosmids with a staff of 10 or so per team organized to carry out the
required tasks in succession. This process is convenient and accumulates data
relatively fast during the early phase of sequencing. The random shotgun strategy,
however, requires challenging efforts prior to sequencing (including the need to
prepare large numbers of sequencing samples, which is largely solved by robotic
operation) and after sequencing (including problems in data assembly and editing
and in filling gaps in the sequence). In its current application, post-sequencing data
processing is the major bottleneck, especially for mammalian genomic DNA, which
is enriched in GC content and repetitive sequences at intervals.

Our laboratory is testing an alternative approach based on the notion that
rather than having a large number of technical staff, each carrying out a few of a
long series of steps, one or two individuals could accomplish megabase equivalents
of sequencing per year (a ten-fold increase over present standards) if 5 kb or so
were sequenced each day in an automation-assisted way, and efficient ways were found to merge the sequence information from the output on different days. Such approaches would be more adaptable to distributive sequencing at many sites, but are dependent on the introduction of some order into the sequencing process.

Ordering can be achieved in a number of ways. Fingerprinting and restriction mapping of clones are the most established, but we are instead attempting to construct the map during the sequencing process itself. This approach, Ordered Shotgun Sequencing (OSS), is achieved by subcloning a large clone like a YAC, systematically sequencing the ends of the subclones, and using overlaps of the sequenced ends to develop a partial map. Subclones (say, with 5 kb inserts) at the edges of contigs are then fully sequenced (by the shotgun method mentioned as above, with one person sequencing 5 kb per day). The results seed the sequencing of additional subclones minimally overlapped across the YAC. Features of the OSS approach include: (a) quickly obtainable single-tract sequencing information is used to construct a map early in the process; (b) facilitation of flexible choices about regions to be sequenced; (c) reduction of the required sequencing redundancy; (d) better discrimination of repetitive sequences within a cloned segment; (e) easy automation, since an auto-sequencer is the only major equipment; and (f) easier project management, with 5 kb sequence units daily. A feasibility study using lambda DNA as the test target suggested that sequencing 12 DNA fragments (about 5 kb each) would be enough to complete a 50 kb project.

In its current stage of development, the implementation of OSS crucially depends on the capacity to sequence efficiently PCR products amplified from lambda or plasmid subclones. By modifying PCR conditions and using a new sequencing enzyme (a Taq polymerase mutant), we have developed a protocol which links PCR directly to cycle sequencing without any intervening step. This procedure thus both expedites the determination of end sequences from the colonies or plaques and facilitates the shotgun sequencing of each 5 kb fragment. Current conditions provide a throughput of two plasmids/week, and with additional robotic
aids, it is conceivable for an investigator to aim at the analysis of a net of 5 kb per day, or the order of 1 Mb of DNA per year.

Several new technologies are designed to help order sequence tracts and to increase the rate of sequencing by several orders of magnitude. They include sequencing by hybridization to oligonucleotides, using mass spectrometry, single molecule sequencing, atomic probe microscopy, etc. These approaches are in an early stage of development, and none has yet been put to practical test. But SBH, for example, can potentially provide a useful adjunct to conventional sequencing strategies. This is because, in principle, the oligonucleotide patterns generated from SBH experiments could help to order random clones, greatly benefiting shotgun strategies.

Other techniques have been developed that can appreciably improve the efficiency of current methods. For example, capillary electrophoresis can expedite the speed of DNA size separation several-fold, and multiplex sequencing schemes could significantly reduce the number of gel runs. But despite their promising features, these methods have idiosyncratic problems and have not yet been widely adopted. Furthermore, they do not alleviate the cost-limiting steps in current practice.

Regardless of whether a revolutionary technology materializes in the longer term, incremental improvements in existing approaches dominate in the short term. Increases in efficiency could be augmented by improvements in current automated sequencing systems. Those will most likely include longer sequence tracts from more channels on a gel; more sophisticated computer software to assemble, edit, and compare data; and new strategy like OSS which can help to make a bridge from mapping to sequencing.

NEW BIOLOGY NEEDS NEW APPROACHES: DNA SEQUENCING BY HYBRIDIZATION AND PROTEIN MAPPING ON DNA BY CROSSLINKING

A. Mirzabekov
Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow.

The programs to sequence the whole human genome and the genomes of other animals, plants, and microorganisms are opening new avenues in modern biology. The need to read billions of nucleotides demands an essential improvement in existing technologies as well as the development of new approaches. At the same time enormous amount of information that will be provided by DNA sequencing requires a search for commensurate, new experimental methods.

Considering these demands, two methods are described, one for DNA sequencing by hybridization to oligonucleotide microchips (SHOM) and another for mapping proteins on genomic DNA by DNA-protein crosslinking.

SHOM is based on the hybridization of the DNA to be sequenced with an enormous set of short oligonucleotides immobilized on a microchip. Identification of the overlapping set of oligomers that form perfect duplexes with the DNA permits reconstruction of the DNA sequence (1). Technology will be described for the production of microchips by immobilization of the presynthesised oligonucleotides on polyacrylamide gel elements of the microchips, A specially devised fluorescent microscope supplied with a CCD-camera and a computer is able to monitor quantitatively the hybridization of fluorescently labeled DNA with immobilized oligomers. The result of SHOM application for model sequencing and for diagnostics of mutations will be presented. Advantages, shortcomings, and further development of SHOM as a simple, fast, and inexpensive method for DNA sequencing and sequence comparison will be discussed.
To reveal the mechanisms of DNA packaging and functioning in transcription, replication, and other processes. hundreds of proteins bound to enormous numbers of DNA segments should be identified and the sites of their interaction. in the DNA and in the protein molecules. should be localized. Our laboratory has been developing procedures of two-dimensional retardation gel electrophoresis to localize such crosslinked sites after covalent bonding Application of these methods has enabled us to identify the position of histones on nucleosomal DNA, to localize the interacting regions within the histones, and to map a number of proteins on regulatory elements of several genes within genomic DNA (2). The use of the methods for large-scale mapping of proteins on genomic DNA will be discussed.

References

CONSTRUCTION OF PRIMERS BY LIGATION OF SHORT
DEGENERATE OLIGOMERS AND RAPID SEQUENCING
WITHOUT CONVENTIONAL CLON-ING
W. Szybalski
McArdle Laboratory, University of Wisconsin, Madison, WI 53706. U.S.A

At present, most of the large genome sequencing projects are based on the
“bottom-up” approach, as applied to the libraries of large DNA fragments cloned in
various vectors. We, on the other hand, favor the “top-down” approaches consisting
of the physical mapping and preparation of contiguous 50-100-kb fragments directly
from the genome, followed by their automated sequencing based on the rapid
assembly of primers by hexamer ligation together with primer walking. Our “top-
down” procedure totally avoids conventional cloning, subcloning and random
sequencing, which are the elements of the present “bottom-up” procedures
[Szybalski, Gene 135 (1993) 279-290]. Physical mapping employs rare-cutting
restriction systems (including Achilles’ heel cleavage [AC] or the RecA-Ac
procedures of Koob et al. [Nucleic Acid Res. 20 (1992) 5831-5836]. The same
methods could be used for in vitro excision of 50-100-kb fragments, ready for
automated sequencing. However, we are developing more efficient methods for in
vivo excision and amplification of 50-100-kb DNA fragments, using the yeast
FRT/ Flp system or phage λ att/Int system. Such fragments, when derived directly
from the Escherichia coli genome, are all arranged in consecutive order, so that 50
specially constructed strains of E. coli would supply 50 end-to-end arranged approx.
100-kb fragments, which will cover the entire approx. 5-Mb E. coli genome. Using
the FRT/ Flp excision system and plasmid replication system, both under tight
transcription control, we succeeded to excise and amplify several 5--100-kb
genomic fragments directly from the host genome [Posfai et al., Nucleic Acid Res.
22 (1994) 2392-2398]. To apply our in vivo excision/amplification system to the
150-Mb Drosophila melanogaster genome, 1500 such consecutive 100-kb
fragments (supplied by 1500 strains) are required to cover the entire genome.
Proportionally more strains would have to be constructed for organisms with larger
 genomes. At present, we are improving the amplification/replication component of our system, as to make it more universal: the excision component, on the other hand, works in a broad variety of prokaryotic and eukaryotic species.

Once the 50-100-kb fragments of DNA are prepared, they will all be sequenced by the SPEL-6 method involving hexamer ligation [Szybalski, Gene 90 (1990) 177-178; Fresenius J. Anal. Chem. 4 (1992) 343] and primer walking. The 18-mer primers are synthesized in only a few minutes from three complementary and contiguous hexamers annealed to the DNA strand to be sequenced when using an over 100-fold excess of hexamers and T4 DNA ligase at room temperature, preferably in the presence of the single-strand-binding (SSB) protein of *E.coli*. These 18-nucleotide (nt) primers are immediately extended by the DNA polymerase, Sequenase 2.0, in the dideoxy sequencing reaction. Very high quality sequencing ladders are obtained for single-stranded DNA or denatured double-strand approx. 50-kb fragments, as exemplified by phage λ DNA. The method is described by Kaczorowski and Szybalski [Analyt. Biochem. 221 (1994) 127-135]. Theoretically, a library of 4096 hexamers is required for SPEL-6 primer walking, but we have shown that less than 1000 hexamers will suffice, since we could use all four nt at position 3 of all hexamers, which reduced the library by a factor of four.

The present SPEL-6 method, when automated and used in conjunction with fluorescent dyes and ultra thin gels, should permit the sequencing of 500 nt per 30 min., i.e., 1kb/hour and 100 kb in less than a week per one sequencing channel. Automation has to include direct gel readout of over 500 nt, analysis of the terminal 50 nt, computerized selection and robotic assembly of 18-mers from three hexamers followed by their template-dependent ligation, sequencing reactions, instantaneous deproteinization, gel loading, electrophoresis, and again a gel readout followed by the next cycle. With 50 channels and all approx. 100-kb genomic fragments available (see above), one could project that automated sequencing of the entire *E.coli* genome should take about one week. Sequencing of both strands and larger genomes would require proportionally more time or more automated sequencing.
equipment. There is little doubt in my mind that automated “top-down” approaches are the key to the efficient and rapid sequencing of large genomes.
FLUORESCENT DIFFERENTIAL DISPLAY: A HIGH THROUGHPUT SYSTEM FOR A LARGE-SCALE SCANNING OF THE DIFFERENTIALLY EXPRESSED GENES

Takashi Ito, Keiji Kito, Naoki Adati, Yuko Mitsui, Hisashi Hagiwara, and Yoshiyuki Sakaki
Human Genome Center, Institute of Medical Science, University of Tokyo,
4-6-1 Shirokanedai, Minato-ku, Tokyo 108, Japan

Introduction

Among tens of thousands different genes in higher eukaryotic genomes, particular subsets are expressed in a temporally and spatially regulated manner to accomplish normal development, growth and differentiation. Identification of such differentially expressed genes are of particular importance as a first step toward the understanding of molecular mechanisms underlying the coordinated regulation of gene expression.

A unique PCR-based technique, termed differential display, has been developed by Liang and Pardee (Science 257967-971, 1992) as a novel approach suitable for such purpose. Basic concept of this approach is to sample a defined subpopulation of transcripts through RT-PCR with arbitrary primers followed by electrophoretic separation of the amplified species as fingerprints on the gel. Simple lane-to-lane comparison among such fingerprints allows one to readily identify differentially expressed species. Unlike subtractive hybridization-based approaches, this method enables simultaneous comparison of more than two samples and identification of transcripts with various behaviors, such as induced, reduced and transiently induced ones.

However, to scan most transcripts in complex organisms, hundreds of primers or primer combinations have to be used. It is thus quite critical for such a large scale analysis to increase the throughput by accelerating each experiment, in
particular, time-consuming steps such as electrophoresis, gel processing and signal detection. Toward this end, we attempted to introduce an automated fluorescent DNA sequencer and developed a novel system, termed “Fluorescent Differential Display (FDD)” (Ito et al, FEBS Letters 351 231-236, 1994 in press). Two simple but robust protocols were developed for FDD. The system allows us to scan nearly 10,000 transcript species a day. They will lay the foundation for a unique high throughput system for expressed sequence scanning, that would be of general use in various biological studies.

**Protocols for FDD**

We tried to run DD on the modified fluorescent DNA sequencer. We routinely use a HITACHI SQ-3000 DNA sequencer which can display gel images just like a fluorescent gel scanner. An important modification of the apparatus is the shortening of the gel length (17cm) to the half of the default one. Such a short gel saves the running time. Bands of DD (up to ~ 600 nt) can be separated within 2.5-3 hrs with satisfactory resolution. Short running time minimizes the damage of gel matrix to enable the re-use of the same gel for the separation of further samples, thereby increasing the throughput considerably. FDD using the primers FITC-TMN (Where M=mixture of A, C and G; N=one of A, C, G and T) and decamers, similar to the original protocol, generated weaker and fewer signals than expected even with various modifications in the reaction conditions. This protocol, termed protocol I, uses arbitrary primers of ~20 nt long and a 3’-anchored oligo-dT primer (Bam TN: FITC-CCCCGATCCT15N, or MluTN: FITC-CGTACGCGT15N, where N=one of A, C or G). Following the first strand synthesis with one of the 3’-anchored oligo-dT primers and reverse transcriptase, the second strand is synthesized during the first cycle of PCR, where a low temperature (37°C) and longer duration for annealing are employed so that even a long arbitrary primer can anneal with multiple targets and initiate the second strand synthesis. In the following PCR cycles, higher annealing temperature (55°C) is used to ensure specific amplification of the primer-tagged cDNA fragments generated.
during the first cycle. A number of bands with enough intensity were readily obtained with much fewer cycles by this method than the protocol close the original one. Another protocol was also established using arbitrary decamers and termed protocol H. (Ito et al FEBS Letters. 351231-236. 1994).

Both protocols generate fingerprints composed of 50-100 (or more) bands in a primer-dependent manner (See Figure). Since we usually load 40 samples per each run, and make three or four runs a day, data on ~ 10,000 transcript species can be obtained a day. The reproducibility of fingerprint pattern is of satisfaction: most (~ 95%) if not all, of the bands seem to be reproducibly amplified. However, faint bands fluctuating from run to run are occasionally observed. Such irreproducible species should not be chosen for further analysis.

To evaluate the sensitivity of FDD, a model experiment was designed, where total RNA from a human glioma cell line U251 “doped” with a known amount of rabbit β-globin-globin mRNA was subjected to the modified differential display protocol using a primer derived from the globin sequence as the upstream primer. We routinely use 50 ng of total RNA for each reaction. The expected band originated from the added globin mRNA was detected even in a sample where the added RNA occupied only 0.00002% of total amounts.

**Molecular cloning of the cDNAs of interest**

The use of fluorescent DNA sequencer allows the rapid identification of the primer combination displaying the band of interesting behavior. Once such a primer pair is identified, the rest of the reaction (~19@) can be used for the preparative isolation of the band. As described below, following the visualization of fingerprint pattern with fluorescent image analyzers or sensitive fluorescent staining, the gel piece was excised and directly used for the re-amplification of the band of interest.

The re-amplified products are in most cases, composed of more than a single species, particularly when non-denaturing gel is used for excision. Thus, we
usually clone the re-amplified products first and, then, identify the correct ones by the digestion of original FDD reaction by appropriate restriction enzymes chosen based on the nucleotide sequences of candidate clones. Following such tests, an oligonucleotide derived from the sequence of the identified clone is synthesized. This specific primer can be used to perform a nested PCR test to prove the identity of the target band and the clone, in which test diluted DD products are used as templates for PCR with the specific and the arbitrary primers. Such PCR would generate the product of expected size and behavior, thereby providing a solid confirmation of the validity of clone selection.

**Examples of FDD analysis**

FDD is currently applied to various projects in our laboratory. For example, during retinoic acid-induced *in vitro* differentiation of a human neuroblastoma cell line SH-SY5Y, many induced and transiently-induced bands were detected, which included a band of the transcript of neurofilament-M, a typical marker of neural differentiation (Kite, K et al unpublished). Another example is the analysis of the transcripts of *Xenopus laevis* early embryo (Adati, N et al manuscript in preparation). Bands of various behaviors, including induced, reduced, stage-specific and unchanged ones, were readily observed, reflecting the dynamic change in gene expression programs during early embryonic development.

We have so far cloned 22 bands showing differentially expressed or polymorphic FDD patterns from various sources. Subsequent Northern blot hybridization and/or RT-PCR analysis using specific primers derived from the identified clones proved the validity of FDD pattern in all the cases tested, which showed the robustness and reliability of our procedures. We attribute the high success rate not only to the careful clone examination step but also to the novel primer design that allows the use of more standard and well-established RT-PCR conditions compared with the unusual ones in the original DD. The protocols
described here will facilitate the identification of differentially expressed as well as polymorphic transcripts to address various biological questions.

Legend to Figure

FDD profiles of the transcripts of the neuroblastoma induced by retinoic acid \textit{in vitro}. Five different sets of primers were used.
Development of Simple Molecular Biology Techniques for Genome Research in Latin America

Sergio D.J. Pena (1, 2) and Andrew J.G. Simpson (3)

1) Departamento de Bioquimica e Imunologia, Instituto de Ciencias Biologicas - UFMG, Avenida Antonio Carios, 6627, 31270-901 Belo Horizonte, MG, Brazil.
2) Nucleo de Genetica Medica de Minas Gerais (GENE), Av. Afonso Pena 3111/9, 30130-909 Belo Horizonte, MG, Brazil.
3) Laboratorio de Biologia Molecular, Centro de Pesquisas “Rene Rachou” - FIOCRUZ, 30190-002 Belo Horizonte, MG, Brasil.

Correspondence to: Dr. S.D.J. Pena, Departamento de Bioquimica e Imunologia, Instituto de Ciencias Biologicas - UFMG, Caixa Posta 486, Belo Horizonte, MG 30161-970 Brazil.
Tel. (55-31) 227-3496; Fax (55-31) 227-3792;
e-mail: spena@vml.lcc.ufmg.br

Introduction

The praxis of genome research is technique-intensive. Novel methodologies often have radical impact by opening whole new areas for investigation. For example, the introduction of the polymerase chain reaction (PCR) represented a veritable revolution in “democratizing” molecular biology, rendering it amenable to practice in the Third World. In the past few years our research group has been dedicated to the development of simple, inexpensive, non-isotopic techniques suitable for genome research in the ill-equipped low-budget laboratories of Latin America. Among the many techniques that we have cultivated are the early use of non-isotopic biotin-labeling of probes (Mederos et al., 1978; Macedo et al. 1979; Pena et al., 1981), utilization of in-gel hybridization (Gontijo and Pena, 1980), employment of chemiluminescence and alkaline phosphatase-conjugated oligonucleotide probes (Pena et al., 1993), improvement of sensitive silver staining techniques for polyacrylamide gels (Santos et al., 1993; Dias Neto et al., 1993a;
Sanguinetti et al., 1994) and betterment of control conditions for PCR (Dias Neto et al., 1993b). We here report a novel simple and rapid PCR-based technique for detection of DNA sequence variation.

Recently there has been extraordinary interest in the development of techniques capable of detecting mutations in DNA sequences. While DNA sequencing is obviously the most sensitive and informative method: it is still too cumbersome for routine use, especially when the DNA fragment of interest is large. For the detection of single base mutations, the alternatives to sequencing consist of techniques dependent on prior knowledge of the exact base alterations (site-specific approaches) or shotgun strategies that ideally should detect any change. The former involve, for instance, allele-specific hybridization with the direct or in the reverse dot blot procedure. The latter group of shotgun techniques is formed by SSCP, DGGE and others. A complex situation is encountered when one wishes to study sequence differences caused by a variable number of mutations. This is, for instance, the case of identity testing using the hypervariable control (D-loop) segment of mitochondrial DNA where unrelated individuals differ in 1 to 15 nucleotides. Here there is no efficient alternative to sequencing -- site-specific methods become extremely cumbersome due to the large number of potentially varying sites in the sequence while shotgun techniques are not informative enough to distinguish different mutations. On the other hand, direct sequence is complicated by the large size of the D-loop fragment and by the fact that sequencing of both strands is necessary to avoid errors.

We here report a highly sensitive and informative technique generally applicable to the detection of single or multiple mutations in any gene sized DNA fragment (Pena et al., 1994a). It consists of submitting a purified DNA fragment to multiple cycles of PCR amplification in the presence of a single oligonucleotide primer specific for one of the extremities of the fragment under conditions of very low stringency. The primer hybridizes with high specificity to its complementary extremity and with low specificity to multiple sites within the fragment, in a sequence-dependent manner. The reaction thus yields a large number of products
that can be resolved by polyacrylamide gel electrophoresis to give rise to a multiband DNA fragment “signature” that reflects the underlying sequence. Changes as small as single base mutations will alter drastically the multiband pattern producing a new signature that is diagnostic of that specific alteration. We have called the technique LSSP-PCR (low-stringency single specific primer PCR). In this manuscript we briefly describe two useful applications of LSSP-PCR: as an identity test in humans and as a tool for strain identification of human papillomavirus.

A New Human Identity Test

The human mitochondrial DNA presents several characteristics that make it useful for human identification and establishing family relationships (e.g., Ginther et al., 1993). First it is haploid, not undergoing recombination, and exhibits matrilineal inheritance. Second, it is highly variable, mostly in the 1,024 base-pair control (D-loop) region, which evolves 5-fold faster than the remainder. Apparently unrelated Caucasian individuals are very rarely identical, differing on average on 7 nucleotides. This makes mtDNA matrilineage-specific. The identity of an individual can thus be tested by comparison with any matrilineal relative. Third, mtDNA is very abundant, being present in up to 10,000 copies per cell. This natural amplification is convenient when dealing with small or degraded samples. The main disadvantages of the study of mitochondrial DNA as compared to nuclear genome markers are that paternity cannot be tested and that it cannot be used to distinguish between members of the same matrilineage. This latter idiosyncrasy limits its use in the identification of criminals. However the study of mtDNA is still extremely useful in the identification of victims of crimes or accidents.

To apply LSSP-PCR to the D-loop (control region) of human mitochondrial DNA, we prepared human DNA samples from blood and amplified by standard procedures the 1,024-bp portion of the mtDNA control region between primers L15996 and H408. The amplification product was then purified by electrophoresis in agarose gel and subjected to a second amplification of 40 cycles at very low
stringency (annealing temperature = 30°C) using several concentrations of a single primer (L15996) and Taq polymerase. The amplification products of PCR amplification were then analyzed in polyacrylamide gels and silver stained. At very high concentrations of both primer μM compared to the normal range of 0, 1-1 μM and Taq polymerase (160 U/ml) a complex pattern consisting of dozens of bands was obtained. Whenever the concentration of either component was decreased there was a sharp drop in the number of visible bands.

Our next step was to ascertain whether the D-loop LSSP-PCR signatures varied in unrelated individuals. For this, we compared more than 40 unrelated individuals and always obtained different multiband signatures. In contrast, 30 mother-child pairs were always identical, as expected from the matrilineal inheritance of mitochondrial DNA. The mother and child pairs show absolutely identical signatures, which are quite different from the father. We conclude that LSSP-PCR is a powerful technique for comparing multiple sequence differences in DNA fragments and that it will provide a useful new tool for identity testing by comparison of mtDNA D-loops (Pena et al., 1994b).

Human Papillomavirus Strain Identification

Human papillomavirus (HPVs) form a taxonomically complex group of organisms responsible both for benign warts and malignant lesions, particularly of the anogenital tract. There are approximately 70 HPV types defined on the basis that each has less than 90% homology in their 11 and E6 genes and long control region with any other type. Each type is in turn heterogeneous consisting of an unknown number of variants with higher levels of homology. An analysis of HPV-16, the most common type associated with genital neoplasias, revealed 4 different variants amongst 301 isolates collected worldwide. Different HPVs give rise to different pathologies and some HPV types such as 16 and 18 are highly associated with anogenital cancer while other such a 6 and 11 are not. The association of particular variants of HPV-16 and 18 to anogenital cancers remains to be established.
The simplicity of LSSP-PCR prompted us, in collaboration with Dr. L.L. Villa from the Ludwig Institute in Sao Paulo, to investigate its application in the area of virus diagnosis as a means of identifying clinical specimens based on sequence variation in the diagnostic PCR product. The aims of the work were to establish conditions for producing LSSP-PCR gene signatures from the L1 gene fragment amplified by the widely used MY09 and MY11 degenerate primers, to ascertain the range of different types that would produce useful gene signatures and to investigate the alterations in gene signatures produced when both different type as well as variants of the same type are compared. With a primer internal to the L1 gene fragment we could successfully perform LSSP-PCR from all specimens. All samples tested produced relatively complex gene signatures consisting of at least two intense bands, in the range of 200 to 400 bp, together with a variable number of the abundant products. The patterns observed were sample dependent and varied between HPV types. Moreover, different gene signatures were obtained for the HPV-16 and -18 prototypes and two HPV-positive cell line, SiHa and HeLa, known to bear variants of HPV-16 and -18 DNA, respectively. This demonstrated the capacity of the methodology to distinguish between variants of the same type, Nevertheless, the signatures derived from the same sample were reproducible and not dependent on template concentration. Altogether, these observations suggested that the L1 gene template used together with the internal primer for LSSP-PCR formed the basis for an HPV identification protocol. A manuscript describing these results is in press (Villa et al., 1994).

**Other Applications**

Although we have presented only two examples, LSSP-PCR has unlimited applications. Among the other uses that we have already made of it are the identification of single base pair mutations in genetic diseases, discovery of polymorphisms in human genes and the classification of strains of *Trypanosoma cruzi*. LSSP-PCR should be of value in essentially all situations where the rapid
detection of variable and unpredictable changes in defined sequences is required. We find it reasonable to suppose that it will be universally applicable.

Acknowledgements

Part of this research was funded by grants-in-aid from FAPEMIG and CNPq. The technique of LSSP-PCR described in the present report is the subject of a patent application.

References


NEW CONCEPTS IN PATHOGENESIS OF SLOW VIRAL INFECTIONS
D. Carleton Gajdusek
National Institute of Health, Bethesda

Slow infections with unconventional viruses ---- Kuru, Creutzfeldt-Jakob disease (CJD), Gerstmann-Strausler syndrome (GSS), Scrapie, bovine spongiform encephalopathy ---- are now recognized as transmissible amyloidoses of brain. As with most amyloidoses, the spontaneous generation of amyloid fibrils is under genetic control. Amyloidogenesis is facilitated by amino-acid-changing point mutations. These infectious amyloids are the causative agents of the transmissible subacute spongiform encephalopathies for which Koch’s postulates for an etiological transmissible replicating agent have been satisfied.

The infectious amyloids of the transmissible brain amyloidoses are generated from a normal host precursor protein by nucleation of confirmational change to a more stable cross-β-pleated lower energy state by tight epitaxic (geometrical) matching of the normal precursor monomer by hydrogen bonding, and perhaps stronger bond formation, with or without the aid of a ligand. This confirmational transition results in spontaneous de novo generation of an infectious molecule without a chain of infection. The CJD-amyloid host precursor protein is specified on chromosome 20 in man; the homologous scrapie-amyloid precursor protein is specified on chromosome 2 in mice. Its normal function is unknown, but it is anchored to the external membrane surface by a conventional inositol-stearic acid anchor and its nucleated transition to an infectious amyloid may occur extracellularly to only a rare molecule. However, this protease resistant infectious amyloid is no longer rapidly turned over and these molecules polymerize into insoluble amyloid fibrils internalized into lysosomes by endocytosis. This occurs particularly in low-density leukocytes and in neurons. Only the long-lived, non-dividing neurons develop pathology after month, or years, of slowly accumulated rising infectivity. How the vacuolating cytotoxicity is produced is unknown.
The 254 amino-acid containing human CJD-amyloid precursor contains 5 copies of an octapeptide in tandem repeats. Families with 2, 4, 5, 6, 7, 8 or 9 additional octapeptide copies develop autosomally dominant familial CJD. Thus, these inserts, as do the amino-acid substituting mutations, increase the stochastic likelihood of spontaneous transition to infectious conformation about a million-fold (the world-wide incidence of sporadic CJD), producing a regular occurrence of familial CJD in autosomal dominant pattern rather than the rare stochastic spontaneous event of sporadic CJD, in which there is no mutation.

Different mutations of the precursor produce different clinicopathological forms of transmissible brain amyloidosis (CJD, GSS, etc.) or variable clinical pictures.
THE PATIENT RECOVERS
E. Grey Dimond, M.D.
University of Missouri-Kansas City, 2501 Holmes Street. Kansas City,
Missouri 64108, U.S.A

The title for my remarks. THE PATIENT RECOVERS, is a straightforward
health bulletin about our host country.

My metaphor is deliberate: to speak of China as a sick patient that after a
protracted illness initiated by a very invasive virus, with multiple therapies applied
by a horde of rapacious “physicians”, a medley of near fatal surgeries and blood-
lettings, excessive narcotics, medical bills disguised as reparations, a sick patient has
now passed through convalescence, rehabilitation, and is walking, even running, to
catch up on time lost.

China was almost felled by that European epidemic called colonizing, and
brought into a near terminal state by the Asian variety of the same. When all hope
was marginal, with parts of the body being passed out for transplantation, a host of
representatives of the pious, and of alien dogma, came and each brought their
version of healing, And demanded that the patient take their dosage with their
incantation, Somehow from all this misery, pain, and debilitation, and almost inspite
of extremely radical measures, the patient has recovered, not only recovered but
seems to have acquired a joie de vivre unique among similarly stricken patients.

Let me leave my simile, my metaphor, and speak of the reality. Here at this
extremely sophisticated conference, a gathering of international talent dealing with
the very outer margins of Humankind’s understanding of what makes Humankind --
the very gene -- this ancient, glorious culture, society, presence called China, is not
only the host but a full participant in this highest form of the scientific era, this
ultimate example of the international, cerebral, 21st century frontier. China is back.
Back not only as the host presenting a culture for our enjoyment that has amazed
the foreigner for centuries -- but back as a fully participating, contributing member of international science.

I have been coming to China for almost a lifetime and in making that statement I prove how inadequate one man’s view can be. My lifetime is but thirteen hundredths of one percent of this culture’s continuity. How little I can understand, how much has rolled on through the millennia, how imperceptible my biopsy. But that is all one man has as his share, From that constrained view. I raise my figurative glass and toast our hosts and congratulate them, each and all. for their recovery from a near fatal illness. Gambei!
MEDICAL GENETIC SERVICES IN HUMAN GENOME RESEARCH

Norio FUJIKI. (Emeritus Professor, Fukui Medical School(1). Medical Advisor. Gene Analysis Laboratory, Tsuruga TOYOBO(2)), with cooperation of Masaru KURIYAMA, (Prof. Dept. Med. (1)), Takeshi AZUMA, (Instructor, Dept. Med. (1)) and Kozo HASHIMOTO, (Head. (2))

Through the clinical applications of human genetic knowledge, early diagnosis, neonatal screening, prenatal diagnosis followed with selective abortion, carrier detection as well as genetic counseling, together with such preventive measures as prohibition of inbreeding, eugenic law, family planning and treatments, have been most effective, as shown in WHO thalassemia eradication strategy.

We have experienced on genetic counseling businesses totaled of over 3000 cases in 3 different areas of Kyoto, Nagoya and Fukui, and made the follow-up studies. After these studies, we have performed consensus survey on heredity, handicapped and new biotechnology among lay peoples in domestic (2000) and international (1500] levels, with the support of WHO and UNESCO, and found many mis-understandings and prejudices, although there were differ-ences in religious, cultural and social structures.

Recently, tremendous developments of human genome research have been applied to our genetic counseling businesses, especially on such intractable neurological diseases as SBMA, MD, and Fra(X), using tri-nucleotide repeat and on genetic susceptibility for common diseases, using RFLPs.

The new bioethical problems as the justification of presymptomatic diagnosis of Huntington disease, and the possibility for discrimination of patients with genetic and even common and psychiatric diseases using predictive markers as well as the availability of improvement of health welfares for handicapped children, has been further discussed,
In order to educate general public for their misunderstandings and prejudices, as shown in our re-searches, we should emphasize more time for medical and postgraduate curriculum of medical genetics and bioethics. Medical genetics is the science itself for the study of individuality and variability, that the DNA sequences of each person are greatly varied, as Human Biodiversity Program has emphasized. Thus such great diversity should be respected, in order to preserve the dignity of mankind.
STUDY ON PHYSICAL MAPPING OF HUMAN CHROMOSOME 17q11-q12 BY CONSTRUCTION OF YAC CONTIGS

Shunsheng Han,'Long Yu,'Jiahui Xia, *Zhu Chen. °Yu Deng, 'Shouyuan Zhao

'Institute of Genetics, Fudan University, 220 Handan Road, Shanghai 200433, P. R. China,
* State Key Laboratory of Molecular Biology, Hunan Medical University, Changsa, Hunan 410078, P. R. China.
°Institute of Hematology, Ruijin Hospital, Shanghai 200030, P.R. China.

Human chromosome 17q11-q12 is a focus of human genome mapping since many oncogenes or antioncogenes were located and cloned in the region, such as ERBB2 neurofibromatosis type I (NF1) gene, retinoic acid receptor, alpha (RARA) gene and breast and ovarian cancer (BRCA) gene, etc. recently several yeast artificial chromosome (YAC) contigs covering NF1, GCSF, EPB3, RARA locus have been constructed. And these contigs are very valuable for mapping chromosome 17q11-12. Presently we have constructed several contigs with 19 probes of 17q11-q12.

In order to get enough DNA markers to finish the physical mapping of the region, a specific DNA probe pool were constructed with Sau 3AI fragments microdissected at 17q11-q12. Twenty single copies were isolated and identified successfully from it with linker and PCR amplification. Eleven of the 20 probes were used to screen Imperial Cancer Research Fund (ICRF) YAC library. And these probes are being sequenced in order to obtain sequence tagged sites to screen CEPH YAC library besides, more specific DNA single copies are being isolated and identified from our probe pool.

One hundred and thirteen YAC clones were picked up by screening ICRF YAC library by hybridization to the filters which were spotted with high density...
YAC clones with 19 DNA Probes located in 17q11-12, of which 90 clones were selected by the 11 new probes from microdissected pool, and the other 37 YAC clones picked up by the 8 probes from ATCC. Among the 113 clones 19 clones were hybridized positive to more than 2 probes. After analyzing the relationship between probes and positive clones, three contigs were obtained at least. They are as follows: a contig covering erythrocyte surface protein band 3 (EPB3) gene region constructed with probe EPB3 and 4 new probes, a contig covering Thyroid hormone receptor alpha 1 gene (THRA) constructed by probe THRA and 2 new probes, a contig covering crystalline beta polypeptide 1 (CRYB 1) region constructed by probe CRYB 1 and 3 new probes. After the YAC clones were analyzed with Alu polymerase chain reaction the three contigs were confirmed. The average size of these YAC is 350kb. So the total length of these YAC contigs is supposed to be about 4Mb.

To bridge the gaps and further confirm the YAC ccontigs, isolation of the end of YAC insert is an important but rate limited step. Several methods isolating the end of the YAC insert have been applied, for example, Alu-vectorette PCR, Bubble PCR and plasmid rescue, Presently we have cloned two ends of two YAC clones, which is from CEPH YAC library. One was verified containing repetitive sequence after hybridization with human genomic DNA. Another is a single copy sequence and has been sequenced, The database has been scanned. There is no similar sequence in it. Screening the ICRF YAC library with the end, two YAC clones have been picked up. The ends from the YAC clones above are being isolated by these methods will greatly help us in walking the gaps and mapping the chromosome 17q11-q12.
INTRODUCTION

The objective of animal genomic analysis programme in China is to identify major genes influencing phenotypic traits of economic importance (mainly quantitative trait loci or QTL), or DNA markers closely linked to them. For the approaching to this objective, currently, only genetic linkage mapping or genomic direct scanning can be considered as an efficient tool in practice, and however, this molecular genetic technology is highly dependant on the resolution of animal genetic map, which means sufficient informative DNA markers have been evenly aligned along the whole genome by linkage test. In the linkage map of pig, over 600 markers have been published, mainly co-contributed by the European PiGMaP groups and USDA group at Clay Center USA (Archibald et al., 1994.; Rohrer et al., 1994), and in chicken linkage map, there are more than 400 markers listed (Crittenden et al., 1993; Burt et al., 1994). DNA markers also have been extensively studied and developed for the assignment of genetic linkage map in other livestock species. These maps are now proving effective in the mapping of quantitative trait loci, in other words, for the detection of reasonable effects of quantitative trait loci, the numbers of informative DNA marker presented in the map are more than enough. The need for more genetic markers will be probably necessary to clone and manipulate QTL in order to understand how the genes behalf on the regulation of complex traits, or select favorable QTL alleles via selection on very closely linked markers (marker-assisted selection or MAS) to produce superior breeding stocks. In China, scientific research involved in genomic analysis of domestic animals includes establishment of reference families, isolation of polymorphous DNA markers, construction of large genomic fragment bank with
bacteria artificial chromosome (BAC), \textit{chromosome in situ} hybridization for physical maps, identification of QTL thorough the candidate gene approach, setting or modification of mathematic models for the computation of linkage test or marker-QTL association or marker-assisted prediction. but so far, these research activities are just only being carried out in pig and chicken.

**GENETIC MAPPING STRATEGY**

The divergent breeds of pig used for the construction of reference families are of two types: commercial breed Yorkshire (well-known for the best prolific performance in Western) and Duroc (well-known for the best growth or lean meat trait in the world) crossing to Erhualian-a Chinese local breed with character of the biggest litter size in China. In the first type of reference families, two Yorkshire boars crossing to four Erhualian sows were used to produce the F1 parental generation, members of which were backcrossed to yield F2 individuals for the linkage mapping, especially for the mapping of major fecund loci. In the second type of informative families, two Duroc boars crossing to six Erhualian sows were used to provide F1 offsprings, but within the F1 generation, the intercrossing were employed to yield F2 individuals, which can supply more information on linkage than a backcross, and therefore these families are good for the study of genetic linkage map and could be used to map more QTL in the same crosses. Resource population in China is being aimed to produce 200 F2 individuals in each family, and approximately 40 productive traits will be measured and recorded for analysis.

In chicken, Beijing Red breed was selected to establish the reference families. F0 birds were derived from A line and C line, which extremely differ in many production performances. F1 parental generation were yielded by mating A line cockerel to C line hens, and then F1 birds were intercrossed to produce F2 progeny for segregation linkage analysis. From three generations of the reference population, over 800 chickens were sampled with at least one ml blood, and eight economic traits and 4 qualitative characters have been recorded. Another reference population, by mating a Chinese local breed Taihe (famous for the meat quality and
Chinese medicine) to commercial broiler breed Peterson, is now being under the development.

Most efforts on the identification of DNA markers are taken to clone unique sequence or expressed fragments named Class I marker, which usually have only two alleles and can be highly informative in a cross between divergent breeds fixed for different alleles. Some mini satellite markers have been isolated and sequenced by screening lambda phage or cosmid libraries with the probe of core sequences published in human or mouse. RAPD (random amplification polymorphism DNA) marker also have been extensively studied in recent years, which employs polymerase chain reaction (PCR) technique coupled with arbitrarily random short primer — usually only ten nucleotides, and however, these markers are dominant and not quite stable in many cases, which means that sometimes they are not following the mode of Mendelian inheritance. In the last year, more emphasis was put on to identify simple tandem repeat loci (or micro satellites) as the markers of choice for genomic scanning, which generally are highly polymorphous, producible and now can be genotyped by semi-automation.

BAC library for the construction of high resolution physical map based upon contigs of overlapping large genomic fragments is much better than conventional yeast artificial chromosome (YAC) library in several parts. The average insertion in BAC vector is able to contain the complete domains of QTL, so the manipulation of QTL for gene transfer will be more convenient. In YAC system, cloning efficiencies are rather low, and chimeras and deletions will cause serious problems when chromosome walking and jumping are performed in “positional” cloning of QTL.

CURRENT STATUS OF GENE MAPPING

Over ten candidate genes have been looked into for the identification of pig major prolific loci, and these genes are more or less involved in pig reproduction regulation system known in physiological level, Up to now, there is no major
prolific loci found. although many polymorphic loci have been detected in the genomic regions of these candidate genes. For typing the genotype of members of informative families from pig reference population in the correlation with the phenotypic traits of interest, DNA markers. some developed in Beijing Agricultural University, mainly taken from the publications or international collaborative laboratories, which mostly will be variable numbers tandem repeat (VNTR) based markers, both those with minisatellite core sequences and microsatellite core sequences, are going to be used to precisely scan one chromosome by one chromosome.

More than 50 restriction fragment length polymorphisms (RFLPs) markers, 100 RAPD markers, nine minisatellites and twenty microsatellites were screened and found to be informative in reference families, but these markers have not been computed to generate linkage groups since not all individuals from the reference families have been genotyped so far. Extensive research also have been carried out on growth performance of pig, mainly by the finding of DNA variation in candidate genes, for instance, growth hormone (GH) gene or insulin-like growth factor (IGF) gene. In practice of pig breeding, genetic variation within population or genetic distance between populations detected by the technique of DNA fingerprint or RAPD-fingerprint were calculated to provide a base for making cross-breeding programme of strategies in conservation of pig species. A PCR-based technique for the detection of carrier of “halothane gene” is now being applied to the breeding programme as a model of “gene diagnosis” in selection, and the mechanism of this technique is based on the identification of point mutation in ryanodine receptor (RYRI) gene, which is demonstrated to be responsible for susceptibility to halothane-induced malignant hyperthermia or porcine stress syndrome (PSS). However, from the results of recent study, halothane gene only can be found in the imported breeds not in Chinese local breeds, although the variations in DNA sequence of RYRI locus have been revealed.

In chicken, the preliminary linkage map currently consists of 32 RFLPs markers, which are dispersed throughout 7 linkage groups. The map covers a
minimum of 223 cM of chicken genome. Over 400 RAPD markers have been identified in reference families and now are being set on the genome by genetic linkage test. More than twenty microsatellites have been sequenced for the primer design. Five genomic clones were mapped by in situ hybridization to provide landmarks for the physical map. In practice of chicken breeding, exploitation of genetic variation or distance among the population by DNA fingerprint or RAPD-fingerprint technique were performed to give the guideline of line-crossing breeding. The fingerprint bands associated with QTL also have been studied, and specific bands significantly linked to age at the first laying, body weight in egg-type breeds were found, and the effects of each band contributes about 15% phenotypic variations, which could be directly used as genetic markers for MAS. The single locus of sex-linked dwarfism and sex-linked feathering were mapped, DNA sequences of which are being accomplished, and the application of these genes information to eliminate heterozygote or select the best alleles in breeding is being performed.

CONCLUSION

The preliminary genetic linkage map of pig and chicken will be achieved by the national collaboration in the next three years, which could be integrated in the map developed by international efforts. The scientific research on gene mapping in cattle and sheep will be carried out soon. Some QTL responsible for 10% variation of production traits could be identified, but the overall application of molecular biotechniques such as “gene diagnosis” for in practice of animal breeding is probably to be on the days of five years later in China.

REFERENCES

STUDY ON MOLECULAR STRUCTURE OF DYSTROPHIN GENE IN CHINESE

Long Yu¹, Ning Wang², Xiaoyun Hau³, Yumei Yang¹, Shengxing Murong², Shouyuan Zhao³

1. Department of Human Medical Genetics.
   Institute of Genetics, Fudan University
   Shanghai. 200433, P. R. China
2. Department of Neurology, Fujian Medical College, Fuzhou,
   30035, P. R. China.
3. Department of Medical Genetics, Sun Yat-sen Medical University

Dystrophin gene is the first gene isolated and cloned by reverse genetics strategy and the largest gene known, covering 2400Kb and about 1/1250 of human genome. An X-linkaged skeletal muscle disorder (Duchenne Muscular Dystrophy, DMD) and an allelic milder form with much slower course are caused by the mutation in this gene. In this study, normal molecular structure and mutation characteristics in Chinese were researched with dystrophin cDNA.

1. The gene structure of 123 cases with DMD/BMD were studied with dystrophin cDNA. 70 cases were detected with deletion mutations, 3 detected with duplications, 1 detected with point-mutation in exon 17. The molecular rearrangement map of dystrophin gene mutation in Chinese has been made according to the extent of the deletion and duplication.

2. Seven partial restriction map of dystrophin gene, which were Hind III, Bgl II, Taq I, Pvu II, Pst I, EcoRV and Xba I, were established with ordering the hybridization patterns from a series of cDNA subprobes produced by overlap-cleavage technique, and the extent of gene mutation. In conclusion, the number of
exons in dystrophin gene should be 79 at least after the map of the hybridization was analyzed in detail.

3. 12 novel RFLPs were revealed by using dystrophin cDNA in this study and comparing the frequency of 22 RFLPs reported previously in Caucasian with Chinese, 3 of them in Chinese were found much higher than the ones in Caucasian, 4 of them in Chinese found much lower than the ones in Caucasian, the difference of both was significant. Furthermore, based on these RFLPs, dosage intensity analysis, 7 partial restriction map, PCR and STR a new strategy was suggested for using to diagnose DMD/BMD heterozygote more rapidly and accurately.

4. Carries states of 140 female from 45 families with DMD/BMD were identified. By analyzing with linkage analysis of RFLPs and gene dosage analysis, 58 defined as normal, and 20 were not determined because of intragenic recombination and less informative RFLPs.

5. The recombination fraction and characterization were studied through the genotype analysis of 512 individuals from 45 families with DMD/BMD.

6. Based on the finding which a 83-year-old male with the deletion from exon 45 to 53 in dystrophin gene still can walk and do some firming, the query about the validity the parathesis that the domain coded by exon 50 and 51 play a key role in the function of dystrophin protein was proposed in the study.

7. Accordant to the sequences of junction fragment produced from RT-PCR amplification for intragenic deletion, the conclusion from our data is correlated with the views of pathogenesis difference in DMD/BMD.
ANALYSIS OF RANDOM AMPLIFIED POLY-MORPHIC DNA
FOR FOUR MINORITIES IN YUNNAN PROVINCE

Wen WANG, Hong LAN, Bing Su, Yaping ZHANG, Li-ming SHI Laboratory of
Cellular and Molecular Evolution, Kunming Institute of Zoology, Chinese Academy
of Sciences, Kunming, 650223. P. R. China

In 1990, two research groups led respectively by Williams and Welsh independently developed a new technique, random amplified polymorphic DNA (RAPD) [1, 2]. In this technique, single oligonucleotides (usually 10bp) of random sequence with 50-70% G+C content are used for polymerase chain reaction (PCR) amplification of DNA. Recently, many studies were published concerning genetic diversity based on RAPD, most of which are on plants [1,2,3,4]. Reports on higher animals and human are much fewer. Now we have conducted a RAPD analysis for four minorities, Lisu, Naxi, Bai, and Tibetan, in Yunnan Province of China in order to add near genetic data in understanding differentiation relationship among these minorities, and simultaneously to assess the applied prospective of RAPD in the study of human genetic diversity.

Eighteen specimens of minorities’ placenta, which were collected from independent pedigrees with the same nationality through at 3 generations, were used to purify genomic DNA. Eighteen random-sequence primers, purchased from Operon Technologies, INC., were used to amplify the purified genomic DNA, out of which 7 primers yielded monomorphic RAPD markers in all individuals, the other 11 varied more or less among individuals.

RAPD appears promising resolution for distinguishing individuals. RAPD markers produced by 18 primers can distinguish every individual experimented, no any two individuals share completely RAPD markers at all loci. We calculated the mean index (1 -F ) of genetic distance between or within nationalities, and found that, 1. Lisu ethnic group has higher genetic variation, whose mean index of genetic distance (1-F ) is 5.465%, 2. the Naxi has lower variation, the average 1 -F is
2.283%, and 3. Bai and Tibetan are relatively distant (1 -F is 5.647%). while Bai and Naxi are closely related (1 -F is 4.421%). It is interesting that the results coincide the geographic distribution and historical origin of the four minorities in the main.

Acknowledgments

This work was partly sponsored by Application and Basic Research Foundation of Yunnan Province. We thank Professor Liu, A. H. and Miss Zhu, C. L. for their technical helps in this study.

References

IDENTIFICATION OF DNA MARKERS LINKED TO BLAST RESISTANCE GENES IN RICE
Zheng KL¹, Qian HR¹, Zhuang JY¹, Lin HX¹, Lu J¹ and Kochert G²

1. Department of Biotechnology, China National Rice Research Institute, Hangzhou, 310006 China
2. Department of Botany, University of Georgia, Athens GA 30602

The identification of RFLP markers closely linked to blast (Pyricularia oryzae) resistance gene is reported for a local Chinese indica rice cultivar Hong Jiao Zhan, which has been widely used as resistance donor for blast resistance in breeding program. Three linked RFLP markers (RG81, RG869 and RZ397) were identified by screening 177 mapped DNA markers (rice genomic clones, rice cDNA clones and oat cDNA clones). Probes were screened by hybridization to Southern blots containing DNA from a pair of near isogenic lines (K80R and K79S) and from the donor and recurrent parents. Linkage association of putatively positive markers with the blast resistance gene was checked using a 143-plant-composed F₂ population segregating for blast resistance. The resistance genotype of each F₂ plants was determined by inoculation of F₃ lines. The blast resistance locus was found to be located on chromosome 12 closely linked to markers RG81 and RG869, and the genetic distances between this locus to RFLP markers were determined.

To fine-map this gene with more DNA markers, RAPD analysis was carried out between isogenic DNA pools from the resistant and susceptible plants of the F₂ population. Polymorphic bands were found between the pools in the products amplified by 6 out of 199 arbitrary primers. Very close linkages between the resistance gene and two RAPD bands (P622-1.3Kb fragment and P286-0.35 Kb fragment) were confirmed after amplification of DNA of all F₂ individuals by these primers. These flanking markers are being converted to PCR markers for marker-assisted breeding.
Half-seed RAPD analysis procedure for blast resistance detection was established. The amplified DNA patterns of the extract from the endosperm half of mature seeds were identical to those of the total DNA from leaves. This procedure will be helpful for the establishment of low cost marker-assisted selection in breeding.
A transgenic mouse strain derived from embryonic stem cells infected with multiple copies of a retroviral vector carries a recessive insertional mutation results in hyperplasia of both embryonic and extra embryonic ectoderm and failure of mesoderm formation in the egg cylinder stage embryo. The mutant phenotype suggests that the gene identified by this insertional mutation plays an important role in the growth control of early embryonic development.

Nodal is a candidate for the mutated gene which encodes a new member of the transforming growth factor-β (TGF-β) superfamily. Whole-mount in situ hybridization of embryos about 7.5 d. p. c. shows that nodal expression coincides with the appearance of the node. The hybridization is highly localized and can be seen as a ring of staining around the node. The expression is not detected by 8.5 d. p. c, coinciding with the disappearance of the node as a distinct structure.

As we know some growth factor mutations may cause tumor, we know nodal recessive mutation may cause cell overproliferate in mouse embryo. We would like to know its function in mature organism, especially in human.

Zoo blot analysis has shown that there are nodal homologous in Drosophila, chicken, rat, vole and human. We are trying to obtain human nodal gene from human genomic library using mouse nodal gene as a probe. We should then sequence this gene and map it to chromosome, and compare it with mouse nodal sequence and its chromosome mapping. Also we would like to compare its polymorphism within Chinese national minorities.
Closure Address on tile South-North Human Genome Conference
Dr. M. Chen, Minister of Public Health, China

Dear Dr. Mayor, Chairman of the International Advisory Committee of the Conference, Director-General of UNESCO,
Dear Dr. Grisolia, Chairman, of the International Organizing Committee, Chairman of UNESCO Human Genome committee,
Dear Dr. Chen, Chairman of the Local Committee.
Ladies and Gentlemen:

The successful opening of the Second South-North Human Genome conference in Beijing, China, is one of the biggest events in the field of medicine and life sciences, at the exact moment of the initiation of Chinese Human Genome Project.

This conference has achieved its great success as expected, under the excellent and fruitful organization and co-ordination of UNESCO, the international and local committees, the Chinese authorities and scientists, as well as all of the participants.

We have been honored to host around two hundreds scientists from more than fifteen countries or regions all over the world, Here I have to mention some of them, Dr. Gajdusek, McKusick, Caskey, among all our distinguished guests.

About forty presentation given by all of you at the Conference have clearly indicated that the human genome project is one of the biggest projects in the history of sciences, and has had a great impact on every respect of medical and life sciences.

China, as a great nation with the largest population in the world, is obviously the biggest beneficial from the human genome project, At the same time, China is also obligated to share the responsibility and effort. China should make its own contribution to the international human genome project.

The Chinese government has given high priority to human genome research by initiating the Chinese Human Genome Project this year, in addition to the Chinese Rice Genome Project. As you know, Dr. Song Jian, State Councilor and Chairman of State Science and Technology Commission, has re-stressed his support to the
human genome project. at a meeting by Chinese scientists and involved authorities on his call immediately before the Conference. Here, on behalf of the Ministry of Public Health of China, I add my firm support to Human Genome Project.

China is still a developing country, and now is open to the world. China and its scientists should work hard, as a member of the international family, to make the Chinese Human Genome Project an integral part of the international effort.

This Conference has evidenced that it has furthered the exchanges between Chinese scientists and those from other countries, and has helped them understand and collaborate with each other.

At last, but not at least, I would like to take the opportunity to express my deep gratitude to:

UNESCO,

The International and Local Committees of the Conference,

All the involved authorities of the Central and Beijing Governments,

Sponsors and all co-sponsors,

All participants and friends,

for their efforts to have made the Conference a real success.

Now, please allow me to announce the closing of the Second South-North Human Genome Conference.
Address at the Closing Meeting of the Second South-North Human Genome Conference
(Nov.9, 1994, Beijing)
Zhang Chongli, Deputy Secretary-General of the Chinese National Commission for UNESCO

Your Excellency Mr. Chen Minzhang,
Distinguished guests and participants.
Ladies and gentlemen.

The three-day meeting is now drawing to a close. During the past three day’s work, by sharing of knowledge and exchange of views and ideas in a very active and lively manner, you have achieved what has bee expected of the goals, and the present conference has turned out to be a great success.

The outcome of this conference will certainly contribute to the advances of the life science and human genome study at international level as well as the initiation and development of the Chinese human genome research programme. For this, allow me to extend, on behalf of the Chinese National Commission for UNESCO and its Chairperson, Ms. Wei Yu, our warmest congratulations.

Distinguished participants,
Ladies and gentlemen,

The great success we have achieved in this conference, in my opinion, is attributable mainly to three major factors. Firstly, the attendance and participation of so many eminent scientists of high calibre in the field of human genome, from both China and abroad, are vitally decisive in achieving the expected results as such scholastically. Needless to say, we are greatly honored by the presence of some of the distinguished scholars of world fame, and to whom my I extend our most heartfelt thanks once again.

Secondly, the success is due to the huge amount of preparatory and organizational work jointly in close co-operation by UNESCO, represented by Ms. Matsui, the international scientific community, represented by Prof. Grisolia and the Chinese
local organizers, especially Beijing University. It would not be fair if I don’t mention Prof. Chen Zhangliang in particular. Without his highly dynamic, efficient and painstaking work, the success of this meeting would not have been possible. So I’d like to express, on behalf of Chinese National Commission for UNESCO, my sincere gratitude to Prof. Chen Zhangliang, Dean of the College of Life Science, Beijing University. My thanks also go to Prof. Gu Xiaocheng, Prof. Gu Hongya, Mr. Wang Yiping and all their colleagues actively engaged in the organizational work.

Last but not least, ladies and gentlemen. the success we have achieved in our conference is, to a large extent, inseparable with the firm support of the Chinese Government. This has been amply demonstrated by the presence in the meeting of H.E. Mr. Song Jian, the State Councillor, himself and also so many ministers and senior officials. The presence of H.E. Mr. Chen Minzhang at both the opening and the closing sessions is also a clear indication of the great interest and concern of the part of the Chinese Government.

Before concluding, I’d like also to represent all the co-sponsors of the Conference, and I am sure that all of you will join me, in expressing our heartly thanks to all the support staff of this meeting for the excellent work they have done for us. And my thanks also go to the international convention Centre, whose good facilities and services are indeed commendable.

Distinguished participants,
Ladies and gentlemen,

The Second South-North Human Genome Conference is now approaching its end. I wish all of you a good time for the rest of your stay in Beijing and a happy journey home.

Thank you
LIST OF PARTICIPANTS
AN. Cheng-cai  
College of Life Sciences, Peking Univ.  
Beijing 100871, China  
Tel: (86)-10-2501862

Ansorge, W  
European Molecular Biology Laboratory  
Meyerhofstr 1  
D-691 17 Heidelberg, Germany  
Tel: (49)-6221-387-355  
Fax: (49)-6221-387-306

Bao, Yi-Ming  
College of Life Sciences, Peking Univ.  
Beijing 100871, China  
Tel: (86)-10-2501847

Bernardi, G  
ICSU’s Committee on Genetic Engineering(COGENE)  
Laboratoire de Recherche en Biologic Moleculaire  
2 Place Jussieu 75005 Paris, FRANCE  
Tel: (33-1)-43-29-58-24  
Fax: (33-1)-46-33-23-05

Cai, R.-S.  
Director, Science Division  
China National Committee for UNESCO

Cai, Tao  
National Laboratory of Cytogenetics  
Human Medical University  
Changsha 410078, China  
Tel: (86-731)-4472093

Caskey T  
Institute for Molecular Genetics  
Baylor College of Medicine  
One Baylor Plaza, T809 Houston  
Texas 77030, USA  
Fax: 1-713-798-7383  
Tel: 1-713-798-4774

Chai, Jian-hua  
Institute of Genetics, Fudan University  
Shanghai 200433, China  
Tel: (86-21)-5492222 ext. 4101

Chen E.Y  
Advanced Center for Genetic Technology  
Applied Biosystems Division, PE  
850 Lincoln Center Drive  
Foster City, CA 94404, USA  
Tel: 1-415-570-66-67  
Fax: 1-415-572-27-43  
E-mail:chenEY@ccmail.apldbio.com

CHEN, Li-rong  
Department of Technical Physics,  
Peking University  
Beijing 100871, China  
Tel:86-10-2501870

Chen, hueih Min  
Department of Biochemistry  
The Hong Kong University of Science and Technology  
Clear Water Bay, Kowloon  
Hong Kong  
Tel: 852-3587294 (Off.)  
852-3405821 (Res.)  
Fax: 852-3581552  
E-mail:BCHMC@USTHK.UST.HK
Chen, Min-zhang  
Minister, Ministry of Public Health of China

Chen, Sai-juan  
Institute of Haemotology.  
Shanghai 2nd Medical University 
Shanghai 200025, China 
Tel: 86-21-4333548

Chen Zhang-liang  
College of Life Sciences, Peking University  
Beijing 100871, China  
Tel: 86-10-2501863

Chen, Zhu  
Institute of Haemotology  
Shanghai 2nd Medical University  
Shanghai 200025, China  
Tel: 86-21-4333548

Chu, Jia-you  
Institute of Medical Biology  
China Academy of Medical Sciences  
Kunming 650107, China 
Tel: (86-871)-8182416

Deane, Bxyan R  
Business Development Manager  
SmithKline Beecham Pharmaceuticals  
SB House, Great West Road  
Brentford, Middlesex TW8 9BD 
Tel: 81-975-3989  
Fax: 81-975-3330

Deng, Han-xiang  
National Medical University  
Changsha 410078, China 
Tel: 86-731-4472093

Deng, Ke-jing  
Institute of Genetics, Fudan University 
Shanghai 200433, China  
Tel: 86-21-5492222 ext. 4103 
Fax: 86-21-5490653

Deng, Yu  
Institute of Genetics, Fudan University  
Shanghai 200433, China  
Tel: 86-21-5492222 ext. 4107  
Fax: 86-21-5490653

Deng, Ke-jing  
Institute of Genetics, Fudan University  
Shanghai 200433, China  
Tel: 86-21-5492222 ext. 4103  
Fax: 86-21-5490653

Deng, Yu  
Institute of Genetics, Fudan University  
Shanghai 200433, China  
Tel: 86-21-5492222 ext. 4107  
Fax: 86-21-5490653

Dellagi, K  
Institute Pasteur de Tunis  
13 Place Pasteur - B.P.74  
1002 Tunis - Belvedere  
Tunisie  
Tel: 216-1-789-608  
Fax: 216-1-791-833

Dimond, E. Grey  
Provost for Health Sciences Emeritus  
University of Missouri  
2501 Holmes street  
Kansas City, Missouri 64108  
USA 
Tel: 1-816-421-5058  
Fax: 1-816-421-5058

Dujon, B  
Institute Pasteur and University P.M. Curie 
Department of Molecular Biology  
25 Rue du Docteur Roux  
F-75724, Paris - CEDEX 15  
FRANCE 
Tel: 33-1-45-68-84-82  
Fax: 33-1-40-61-30-60  
E-mail:bdujou@mendel.sis.pasteur.fr
Hu, Mei-hao  
College of Life Sciences, Peking University  
Beijing 100871, China  
Tel: 86-10-2501859

Kuang, Da-ren  
Shanghai Institute of Cell Biology  
China Academy of Science  
Shanghai 200031, China  
Tel: 86-21-4315030  
Fax: 86-21-4331090

Kuang, Shao-qing  
Institute of Haemotology  
Shanghai 2nd Medical University  
Shanghai 200025, China  
Tel: 86-21-433548

Lau, C.H.  
Department of Biochemistry and Microbiology  
University Pertanian Malaysia  
43400 Serdang, Selangor, Malaysia  
Fax: 60-3-9486646  
Tel: 60-3-9486101-3586  
E-mail: upm@cgnet.com

Li, Ning  
National Laboratory of Agro-Biotechnology  
Beijing Agricultural University  
Beijing 100094, China  
Tel: 86-10-2582244 ext. 701  
Fax: 86-10-2582332

Liu, Hong-ji  
Waters Chromatography’ Division  
USA

Liu, Pu  
National Centre for Human Genome Research, NIH  
NCHGR,NIH,49/3A18  
9000 Rochvillepike  
Bethesda, MD 20892 USA  
Tel: 301-402-2529  
Fax: 301-402-4929

Liu, Yong-hui  
Director  
China National Center for Biotechnology Development  
Beijing, China
Matsui, S  
Program Specialist in Biology  
Basic Science Division  
UNESCO  
P.O. Box 30709  
Nairobi. KENYA  
Tel: 254-2-630-743  
Fax: 254-2-631-499

Mayor, F.  
Director-General of UNESCO  
Penna, S.D.J.  
Latin American Human Genome Program  
Universidade Federal de Minas Gerais  
AV. Afonso Pena 3111/9  
Belo Horizonte, 30130-909  
Brazil  
Tel: 55-31-227-3496  
Fax: 55-31-227-3972

McKusick, V.A  
Johns Hopkins Hospital  
Bialock Building, Rm 1007  
600 N. Wolfe Street  
Baltimore - Maryland 21205  
USA  
Tel: 1-410-95-56-641  
Fax: 1-410-95-54-999

Mirzabekov, A  
Argonne National Laboratory  
Center for Mechanistic Biology & Biotechnology  
9700 South Cass avenue  
Argonne Illinois 60394-4833  
USA  
Fax: 1-708-252-3387

Monaco, Tony  
Imperial Cancer Research Fund  
University of Oxford  
John Radcliffe Hospital  
Headington -Oxford OX3 9DU  
Fax: 44-865-222-431

Ole-Moi Yoi, O.K.  
Biochemistry/Molecular Biology Laboratory  
International Laboratory for Research on Animal Diseases (ILRAD)  
Rynditch, A.  
Institute of Molecular Biology and Genetics  
Ukrainian Academy of Sciences  
150 Zabolotnogo str.
252627 Kiev
UKRAINE
Fax: 7-044-266-3498

Sakaki, Y
Human Genome Center
Institute of Medical Science
University of Tokyo
40601 Shiroganedai, Minato-ku
Tokyo 108, Japan
Tel: 81-3-5449-5662
Fax: 81-3-5449-5445
E-mail: sakaki@hgc.ims.u-tokyo.ac.jp

Schlesinger, D
Washington University School of Medicine
at Washington University Medical Center
Center for Genetics in Medicine
Campus Box 8230, 660 South Euclid Avenue
St. Louis, MO 63110
USA
Fax: 1-314-362-3203
Tel: 1-314-362-2744
E-mail: advids@vugenmail.wustl.edu

Shen, Yan
Institute of Basic Medicine
China Academy of Medical Sciences
Beijing 100005, China
Tel: 86-10-513446 ext. 286
Fax: 86-10-5240529

Sun, Chong-rong
Department of Biochemistry
Fudan University
Shanghai 200433, China
Tel: 86-21-5490323

Sun, Hui
PCTB 805
Department of Molecular Biology and Genetics
The Johns Hopkins University School of Medicine
725 N. Wolfe St.
Baltimore, MD 21205
USA
Tel: 1-410-955-4678
Fax: 1-410-614-0827
E-mail: hsun@welchknk.welch.jhu.edu

Sun, Kai-ai
Laboratory of Molecular Genetics
Medical University of China
Shenyang 110001, China
Tel: 86-24-3863731 ext. 2546
Fax: 86-24-3875539

Szybalki, W
McArdle Laboratory for Cancer Research
University of Wisconsin-Madison
1400 University Avenue
Madison, WI 53706
USA
Fax: 1-608-262-2824
Tel: 1-608-262-1259
E-mail: szybalsk@macc.wise.edu

Tan, Chia-cheng
Institute of Genetics
Fudan University
Shanghai 200433, China
Tel: 86-21-5490323

163
Tsui, L.-C  
Department of Genetics  
Hospital for Sick Children  
555 University Avenue  
Toronto, Ontario N5G 1X8  
CANADA  
Tel: 1-416-813-6015  
Fax: 1-416-813-4931

Wang, Qinghua  
College of Life Sciences  
Peking University  
Beijing 100871, China  
Tel: 86-10-2501857

Ulfendahl, Per Jonhan  
Pharmacia Biotech (China) Ltd.  
Hong Kong

Ward, D  
Department of Human Genetics  
Yale University Medical School  
333 Cedar Street (1-48)  
New Haven, CT 06510  
USA  
Fax: 203-785-7023

Wang, Qing-hua  
College of Life Sciences  
Peking University  
Beijing 100871, China  
Tel: 86-10-2501857

Wei, Yi  
College of Life Sciences  
Peking University  
Beijing 100871, China  
Tel: 86-10-2501840

Wein, Yu  
Chairwoman.  
Chinese National Commission for UNESCO

Wang, Wen  
Kunming Zoology Institute  
China Academy of Sciences  
Kunming 650223, China  
Tel: 86-871-5140390  
Fax: 86-871-5151823

Wei, Yu  
Chairwoman.  
Chinese National Commission for UNESCO

Wu, Guan-yun  
institute of Basic Medicine  
China Academy of Medical Sciences
Wu, Min
Institute of Oncology
China Academy of Medical Sciences
Beijing 100730, China
Tel: 86-10-77150S8

Wu, Shu-qing
President, Peking University
College of Life Sciences
Peking University
Beijing 100871, China
Tel: 86-10-2501848

Xie, Ming
College of Life Sciences
Peking University
Beijing 100871, China
Tel: 86-10-2501862

Xu, Heng-ping
College of Life Sciences
Peking University
Beijing 100871, China
Tel: 86-10-2501840

Xue, You-fang
College of Life Sciences
Peking University
Beijing 100871, China
Tel: 86-10-2501858

Xu, Zhi-hong
Vice President Chinese Academy of Sciences
(Academia Sinica)

Yan, Yong-bin
Institute of Genetics
Academia Sinica

Yu, Long
Institute of Genetics
Fudan University
Shanghai 200433, China
Tel: 86-21-5492222 ext. 4107
Fax: 86-21-5490653

Yu, Zhi-hua
College of Life Sciences
Peking University
Beijing 100871, China
Tel: 86-10-2501847

Yuan, Jian-gang
Institute of Basic Medicine
China Academy of Medical Sciences
Beijing 100730, China
Tel: 86-10-5124876

Yuan, K.-W.
First Secretary,
Chinese Permanent Delegation to UNESCO

Yuan, Yong-bing
College of Life Sciences
Peking University
Beijing 100871, China
Tel: 86-10-2501840
Zeng, Jic  
Group of Genetic Engineering Technology Centre  
Institute of Microbiology  
Academia Sinica  
Tel: 2553521

Zhang, C.-L.  
Deputy Secretary-General.  
Chinese National Commission for UNESCO

Zhang, Cun-hao  
Director, National Natural Science Foundation of China

Zhang, Jian-hua  
College of Life Sciences  
Peking University  
Beijing 100871, China  
Tel: 86-10-2501840

Zhang, Jin-yan  
Silicon Graphics Computer Systems  
Beijing Liaison Office  
Suite 1401, Beijing Fourtune Bidg.  
No.5, Dong San Huan Bei Lu  
Beijing 100004, China  
Tel: 86-10-501-4185 ext.88  
Fax: 86-10-501-4184

Zhang, Si-zhong  
Laboratory of Medical Genetics  
China Western Medical University  
Chengdu 610041, China  
Tel: 86-28-5581130 ext. 2608  
Fax: 86-28-583252

Zhang, Tong  
Institute of Haemotology  
Shanghai Ruijin Hospital  
Shanghai 200025, China  
Tel: 86-21-4370045 ext. 5157

Zharov, V.  
Director, Basic Science Division, UNESCO

Zheng, Jun-hua  
Laboratory of Pharmacology  
Beijing Medical University  
Beijing 100082, China  
Tel: 86-10-2091516  
Fax: 86-10-2015584

Zheng, Kang-le  
Department of Biotechnology  
China National Rice Research Institute  
Hangzhou 310006, China

Zhou, Xun-lei  
College of Life Sciences  
Peking University  
Beijing 100871, China  
Tel: 86-10-2501858

Zhou, Zeng-quan  
College of Life Sciences  
Peking University  
Beijing 100871, China  
Tel: 86-10-2501851

Zhu, Li-huang  
Institute of Genetics  
China Academy of Sciences  
Beijing 100101, China  
Tel: 86-10-2501857

166
Zhu, Li-lan  
Vice Chairwoman  
State Science and Technology Commission  
of China

Zhu, Yu-xian  
College of Life Sciences  
Peking University  
Beijing 100871, China  
Tel: 86-10-2501862