Gamma Field Symposia

Number 32

MOLECULAR EVOLUTION AND MUTATION

1993

INSTITUTE OF RADIATION BREEDING

NIAR MAFF

Ohmiya-machi, Naka-gun, Ibaraki-ken
Japan
MOLECULAR EVOLUTION
AND
MUTATION

Report of Symposium
held on
July 15-16, 1993

Institute of Radiation Breeding
NIAR MAFF

Ohmiya-machi, Naka-gun, Ibaraki-ken 319-22
Japan
The recturers (front row) and staff of the Symposium Committee

The audiences in the Symposium hall

Study tour to the Gamma Field. Around the original pear tree from which "Gold Nijisseiki" was detected.
# List of Participants

(32th GF Symposium)

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution/Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABE, T.</td>
<td>Institute of Radiation Breeding, NIAR</td>
</tr>
<tr>
<td>ADACHI, H.</td>
<td>School of Agriculture, Ibaraki University</td>
</tr>
<tr>
<td>AMAGAI, M.</td>
<td>Tochigi Prefectural Agricultural Experiment Station</td>
</tr>
<tr>
<td>AMANO, S.</td>
<td>Allergen-Free Technology Laboratories, Inc.</td>
</tr>
<tr>
<td>AOKI, C.</td>
<td>School of Agriculture, Nagoya University</td>
</tr>
<tr>
<td>AZUHATA, H.</td>
<td>Tohoku Seed Co., Ltd.</td>
</tr>
<tr>
<td>AZUMA, S.</td>
<td>Niigata Agricultural Experiment Station</td>
</tr>
<tr>
<td>DOMON, E.</td>
<td>Shikoku National Agricultural Experiment Station</td>
</tr>
<tr>
<td>FUJIMAKI, H.</td>
<td>National Institute of Agrobiological Resources</td>
</tr>
<tr>
<td>FUJIMOTO, F.</td>
<td>Faculty of Agriculture, Gifu University</td>
</tr>
<tr>
<td>FUJITA, H.</td>
<td>Matsumoto Branch, National Institute of Sericultural and Entomological Science</td>
</tr>
<tr>
<td>FUJISAWA, A.</td>
<td>Dainippon Ink and Chemicals, Inc.</td>
</tr>
<tr>
<td>FUKUOKA, S.</td>
<td>National Institute of Agrobiological Resources</td>
</tr>
<tr>
<td>GAMO, T.</td>
<td>National Institute of Agrobiological Resources</td>
</tr>
<tr>
<td>HANDA, T.</td>
<td>Kansai Breeding Office, Forest Tree Breeding Institute</td>
</tr>
<tr>
<td>HARA, H.</td>
<td>School of Agriculture, Ibaraki University</td>
</tr>
<tr>
<td>HASEGAWA, H.</td>
<td>Research Institute for Advanced Science and Technology, University of Osaka Prefecture</td>
</tr>
<tr>
<td>HATTORI, K.</td>
<td>School of Agriculture, Nagoya University</td>
</tr>
<tr>
<td>HIRAMATSU, H.</td>
<td>Institute of Agriculture and Forestry, University of Tsukuba</td>
</tr>
<tr>
<td>HIRANO, H.</td>
<td>National Institute of Genetics</td>
</tr>
<tr>
<td>HIRANO, M.</td>
<td>Mie Agricultural Research Center</td>
</tr>
<tr>
<td>HIRASAWA, H.</td>
<td>National Agriculture Research Center</td>
</tr>
<tr>
<td>HONDA, Y.</td>
<td>Saitama Tea Experiment Station</td>
</tr>
<tr>
<td>HOSHI, H.</td>
<td>Forest Tree Breeding Institute</td>
</tr>
<tr>
<td>HOSOKAWA, N.</td>
<td>Agricultural Technology-Institute of Nagano Farmer's Federation</td>
</tr>
<tr>
<td>HYAKUTAKE, H.</td>
<td>The Institute of Physical and Chemical Research</td>
</tr>
<tr>
<td>IIDA, S.</td>
<td>Institute of Radiation Breeding, NIAR</td>
</tr>
<tr>
<td>IKENAGA, H.</td>
<td>National Food Research Institute</td>
</tr>
<tr>
<td>IKETANI, H.</td>
<td>Fruit Tree Research Station</td>
</tr>
<tr>
<td>IKUTA, E.</td>
<td>School of Agriculture, Nagoya University</td>
</tr>
<tr>
<td>IMAI, Y.</td>
<td>School of Agriculture, Nagoya University</td>
</tr>
<tr>
<td>INOUE, E.</td>
<td>School of Agriculture, Ibaraki University</td>
</tr>
<tr>
<td>INUKAI, T.</td>
<td>Faculty of Agriculture, Hokkaido University</td>
</tr>
<tr>
<td>ITO, J.</td>
<td>Tottori Agricultural Experiment Station</td>
</tr>
<tr>
<td>ITOH, R.</td>
<td>Forest Tree Breeding Institute</td>
</tr>
<tr>
<td>KAGEYAMA, C.</td>
<td>Shizuoka Prefectural Citrus Experiment Station</td>
</tr>
<tr>
<td>KAMIO, M.</td>
<td>Agricultural Technology-Institute of Nagano Farmer's Federation</td>
</tr>
<tr>
<td>KANEKO, Y.</td>
<td>Faculty of Agriculture, Utsunomiya University</td>
</tr>
</tbody>
</table>
KASUMI, M. Plant Biotechnology Institute, Ibaraki Agricultural Center
KATO, A. National Glassland Research Institute
KATO, H. National Agriculture Research Center
KATO, T. Tochigi Branch, Tochigi Prefectural Agricultural Experiment Station
KATSUMATA, K. Meiji Seika Kaisha, Ltd., Pharmaceutical Research Center
KATSUWA, M. National Institute of Agrobotanical Resources
KAWAI, T. Sanwa-Soyaku Co., Ltd.
KAWAIDE, O. Tohoku Seed Co., Ltd.
KIRIHARA, T. Plant Biotechnology Institute, Ibaraki Agricultural Center
KITAMURA, K. Forestry and Forest Products Research Institute
KITAURA, T. Agricultural Research Institute of Kanagawa Prefecture
KOJIMA, H. Faculty of Agriculture, Tamagawa University
KOK, H. Iwate Biotechnology Institute
KONDO, T. Forest Tree Breeding Institute
KOSUGIYAMA, M. School of Agriculture, Ibaraki University
KOTSUBUKI, K. Fruit Tree Research Station
KUBOYAMA, T. Faculty of Agriculture, The University of Tokyo
KURUDA, S. National Institute of Agrobotanical Resources
KUSABA, M. Institute of Radiation Breeding, NIAR
MANABE, T. Faculty of Agriculture, Meiji University
MASUDA, T. Institute of Radiation Breeding, NIAR
MATSUMOTO, S. National Research Institute of Vegetables, Ornamental Plants and Tea
MATSUMURA, K. Hitachi, Ltd., Instrument Division, Scientific Instrument System Cluster CSC
MATSUMURA, S. Tohoku Seed Co., Ltd.
MATSUZAWA, Y. Faculty of Agriculture, Utsunomiya University
MITSU, K. Sumitomo Metal Industries, Ltd.
MIYAHARA, K. Institute of Radiation Breeding, NIAR
MIYAMOTO, M. Plant Biotechnology Institute, Ibaraki Agricultural Center
MIYAZAKI, N. Tochigi Prefectural Agricultural Experiment Station
MURAMATSU, N. Faculty of Agriculture, Utsunomiya University
NAGATOMI, S. Institute of Radiation Breeding, NIAR
NAKAGAWA, I. Agriculture, Forestry and Fisheries Research Council
NAKAJIMA, K. National Institute of Agrobotical Resources
NAKAJIMA, Y. Mountaious Region Experiment Farm, Aichi Agricultural Research Center
NAKAMURA, I. Iwate Biotechnology Research Center
NAKAMURA, K. School of Agriculture, Nagoya University
NAKAMURA, T. Rice Breeding Research Laboratories
NAKANISHI, A. National Agricultural Research Center
NAKASHI, T. Nagasaki Agricultural & Forestry Experiment Station
NAKASONE, S. Institute of Radiation Breeding, NIAR
NAKAYAMA, H. National Institute of Agrobotical Resources
NAKAYAMA, S. National Institute of Agrobotical Resources
NAKAZAKI, T. Faculty of Agriculture, Kyoto University
NAMAI, H. Institute of Agriculture and Forestry, University of Tsukuba
NIKURA, S. 
Tohoku Seed Co., Ltd.

NISHIKAWA, Y. 
National Agriculture Research Center

NISHIO, T. 
Institute of Radiation Breeding, NIAR

NISIMURA, S. 
National Institute of Agrobiological Resources

NISHI, M. 
School of Agriculture, Ibaraki University

NOJIRI, C. 
Allergen-Free Technology Laboratories, Inc.

NOMURA, K. 
Institute of Agriculture and Forestry, University of Tsukuba

NUMAGUCHI, K. 
National Agriculture Research Center

OBA, S. 
Faculty of Agriculture, Gifu University

OCHIAI, E. 
Tochigi Prefectural Agricultural Experiment Station

OHTA, T. 
National Institute of Genetics

OHTSUBO, E. 
Institute of the Molecular & Cellular Biosciences, The University of Tokyo

OHYAMA, K. 
Food & Agriculture Research & Development Association

OKADA, K. 
Sanwa-Shoyaku Co., Ltd.

OKADA, M. 
Kochi Agricultural Research Center

OKAMURA, M. 
Forest Tree Breeding Institute

OKUMOTO, Y. 
Faculty of Agriculture, Kyoto University

OTSUKA, M. 
Tochigi Branch, Tochigi Prefectural Agricultural Experiment Station

SAKUMA, F. 
Plant Biotechnology Institute, Ibaraki Agricultural Center

SANADA, T. 
Fruit Tree Research Station

SANO, Y. 
National Institute of Genetics

SASAHERA, T. 
Faculty of Agriculture, Yamagata University

SASAKI, H. 
Q. P. Corporation

SATO, Y. 
National Research Institute of Vegetables, Ornamental Plants and Tea

SATO, Y. 
Yamagata Prefectural Horticultural Experiment Station

SAYOH, T. 
Kennan Branch, Iwate Prefectural Agricultural Experiment Station

SEKIGUCHI, F. 
Faculty of Science, Japan Women's University

SHIMADA, T. 
Allergen-Free Technology Laboratories, Inc.

SHIMANO, S. 
School of Agriculture, Nagoya University

SHIMIZU, A. 
National Research Institute of Vegetables, Ornamental Plants and Tea

SHIMIZU, A. 
Rice Breeding Research Laboratory

SOYOME, T. 
Tochigi Branch, Tochigi Prefectural Agricultural Experiment Station

SUDA, H. 
Tokyo Metropolitan Isotope Research Center

SUDO, M. 
National Agriculture Research Center

SUGAWARA, K. 
Hokkaido Green-Bio Institute

SUMIYOSHI, S. 
National Agriculture Research Center

TADA, T. 
Iwate Biotechnology Institute

TAKAHATA, N. 
School of Life Science, The Graduate University for Advanced Studies

TAKAMURE, I. 
Faculty of Agriculture, Hokkaido University

TAKANO, T. 
Institute of Radiation Breeding, NIAR

TAKATSU, Y. 
Plant Biotechnology Institute, Ibaraki Agricultural Center

TAKEDA, G. 
Faculty of Agriculture, The University of Tokyo

TAKETORI, A. 
Niigata Agricultural Experiment Station

TAMURA, T. 
National institute of Sericultural and Entomological Science
TANAKA, J.  
TANIGUCHI, T.  
TANO, S.  
TERASHI, M.  
TOHNOOKA, T.  
TOMARU, N.  
TOMIYAMA, M.  
TOMOOKA, N.  
Tsuchiya, S.  
Tsumura, Y.  
Tsuyuki, Y.  
Uehara, H.  
Ukai, Y.  
Urushibara, S.  
Wada, T.  
Yahagi, H.  
Yamaguchi, I.  
Yamaguchi, K.  
Yamaguchi, K.  
Yamaguchi, Y.  
Yamamoto, K.  
Yamamoto, T.  
Yamashita, A.  
Yashiro, K.  
Yomoda, A.  
Yoshimura, K.  
Yoshioka, T.  
Zhu, G.  

National Institute of Agrobiological Resources  
Forest Tree Breeding Institute  
Center for Basic Res., Japan Atomic Energy Res. Institute  
Faculty of Agriculture, Tokyo University  
Kyushu National Agricultural Experiment Station  
Institute of Agriculture and Forestry, University of Tsukuba  
National Institute of Agrobiological Resources  
National Institute of Agrobiological Resources  
Fruit Tree Research Station  
Forestry and Forest Products Research Institute  
Faculty of Agriculture, Tamagawa University  
Institute of Agriculture and Forestry, University of Tsukuba  
Faculty of Agriculture, The University of Tokyo  
Iwate Agricultural Experiment Station  
Allergen-Free Technology Laboratories, Inc.  
Ibaraki Prefectural Livestock Experiment Station  
National Agriculture Research Center  
Forest Tree Breeding Institute  
Miyazaki Agricultural Experiment Station  
National Agriculture Research Center  
Kanagawa Horticultural Experiment Station  
Japan Tobacco Inc., Plant Breeding & Genetics Res. Lab.  
Faculty of Agriculture, Tamagawa University  
Institute of Agriculture and Forestry, University of Tsukuba  
Rice Breeding Research Laboratories  
Forestry and Forest Products Research Institute  
Institute of Radiation Breeding, NIAR  
National Agriculture Research Center
FOREWORD

Gene structure has been altered by spontaneous mutation and repeated changes in the gene structure have resulted in the creation of living things. People have developed many valuable crops and livestock utilizing mutation as materials.

Meanwhile, induced mutation has a great potential in crop breeding especially in giving a new character to crop species. For more effective utilization of induced mutation, it would be significant to discuss the mutation with special reference to the evolution of crop species.

We wish to express our sincere thanks to the speakers of special and general lectures, the chairpersons and those who contributed to the success of this symposium.

The Symposium Committee

Tetsuya Abe, Chairperson
Tetsuo Masuda
Shigeki Nagatomi
Kousuke Nakajima
Takeshi Nishio
Yoshio Sano
Genkichi Takeda
Takatoshi Tanisaka
Shichiro Tsuchiya
Yasuo Ukai
PROGRAM

Opening address : T. Abe
Congratulatory address : H. Fujimaki

Special lecture
Chairperson : F. Fujimoto
The nearly neutral theory of molecular evolution and its testing
by DNA sequence analysis .................................................. T. Qhta

Session I
Chairperson : K. Hattori
Genetic variation and gene regulation at the $\alpha\alpha$ locus of Rice .......... H. Hirano

Session II
Chairperson : T. Kondo
Molecular evolution of conifers ........................................... Y. Tumura

Session III
Chairperson : T. Gamou
RFLP patterns detected by the transposon, mag and k1.4, in
silkworm races .............................................................. T. Tamura

Session IV
Chairperson : S. Tsuchiya
Chloroplast DNA diversity in Pyrus and related genera................. H. Iketani

Session V
Chairperson : K. Nakajima
A simple method to classify rice strains with AA genome and to infer their relationships
by identification of transposable elements at various loci .................. E. Ohtsubo

Session VI
Chairperson : Y. Sano
Sexuality and accumulation of deleterious mutations in humans .......... N. Takahata

Session VII
Chairperson : Y. Ukai
General discussion

Closing address : G. Takeda
CONTENTS

T. OHTA
The nearly neutral theory of molecular evolution and
its testing by DNA sequence analysis .......................... 1

H. HIRANO
Genetic variation and gene regulation at the aux locus
of Rice ....................................................................... 19

Y. TSUMURA and K. OHBA
Molecular evolution of conifers ................................. 35

T. TAMURA, T. KANDA, K. YUKUIRO, Y. YASUKOUCHI, N. HINOMOTO, K. SHIMIZU, E. KOSEGAWA and T. OKAJIMA
RFLP patterns detected by the transposon, mag and
k1.4, in silkworm races ............................................. 47

H. IKEJANI
Chloroplast DNA diversity in Pyrus and related genera ...... 63

E. OHTSUBO, K. MOCHIZUKI, T. TENZEN and H. OHTSUBO
A simple method to classify rice strains with AA genome
and to infer their relationships by identification of
transposable elements at various loci ........................... 71

N. TAKAHATA
Sexuality and accumulation of deleterious mutations in
humans ........................................................................ 85

General discussion (in Japanese) ........................................ 95
THE NEARLY NEUTRAL THEORY OF MOLECULAR EVOLUTION
AND ITS TESTING BY DNA SEQUENCE ANALYSIS

Tomoko Ohta
National Institute of Genetics
Yata 1111, Mishima Shizuoka-ken, 411

Introduction

For a long time the study of evolution has been based on morphology; the long neck of a giraffe, the human brain, a bird’s wing, and so on. Morphological change in evolution is explained by DARWIN’s theory of natural selection, but this theory is largely qualitative rather than quantitative. Population genetics started more than half a century ago as an attempt to understand evolutionary change quantitatively, by combininig DARWIN’s theory and Mendelian genetics. Because evolution must take place in all individuals of a species, the change of gene frequency in the population has been analyzed. However, so long as the facts of evolution are based on morphological traits, evolutionary change is very difficult to connect with gene frequency change except in relatively few circumstances.

The remarkable progress of molecular biology has made it possible to apply population genetics theory to real data. We now know that genetic information is stored in linear sequences of DNA which are stably transmitted from generation to generation, and we can compare the linear sequences of DNA and amino acids among species. It is also possible to compare secondary and tertiary structures of proteins and nucleic acids from various sources.

Because of such progress, some aspects of traditional neo-Darwinism are beginning to need revision. The first step in such a revision is the neutral mutation–random drift hypothesis put forward by KIMURA (1968). This theory states that most evolutionary changes at the molecular level are caused by random genetic drift of selectively neutral or nearly neutral mutations rather than by natural selection. Because this theory was contrary to the neo-Darwinian view at that time, it provoked much controversy.

The second step in the revision is to clarify the interaction of natural selection and random drift at the molecular level. Natural selection cannot be so simple as to be “all or nothing”. There are numerous types of mutations, whose behavior is influenced by
both selection and random drift. In this article, theoretical studies of such "nearly neutral" mutations are reviewed, together with some recent findings on DNA sequence variation. The nearly neutral theory will be tested by DNA sequence analysis.

Rate of Molecular Evolution

Since the advent of rapid DNA sequencing techniques, data on the primary structure of genes are accumulating amazingly fast, and now statistical studies of DNA sequences are quite popular. However, only 20 or so years ago, most of the available data on molecular evolution were in the form of amino acid sequences (Lewontin 1974; Ayala 1976). Hemoglobin $\alpha$ of mammals consists of 141 amino acids, and it is one of the best-studied molecules. If one compares human hemoglobin $\alpha$ with that of the gorilla, all amino acids are identical except one, but 18 amino acids differ between human and horse. Such data on sequence divergence faithfully reflect the phylogenetic relationship. Since we know the approximate time of divergence of mammalian species, it is possible to estimate the rate of amino acid substitution. From a comparison of various species the rate of amino acid substitution of hemoglobins $\alpha$ and $\beta$ is about $10^{-9}$ per amino acid site per year. It is quite impressive to find that almost the same value is obtained from any two species. Especially noteworthy is the fact that the rate seems to be almost the same in the line to living fossils whose morphological characters have hardly changed for tens of millions of years, and in the line leading to humans.

The apparent uniformity of evolutionary rates, as compared with phenotypic evolution, is a most remarkable characteristic of protein evolution (Zuckerkandl and Pauling 1965). By applying similar analysis to cytochrome $c$ data, the rate seems again to be almost uniform for diverse organisms including plants, fungi, and mammals. However, the rate is much lower in cytochrome $c$ than in hemoglobins. This is thought to be caused by stronger structural constraints on cytochrome $c$ than on hemoglobins. This is another characteristic of molecular evolution, i.e. the stronger the constraint on the molecule, the lower is its rate of evolution. For some examples of the rates of protein evolution, see (Kimura 1983). It is now well known that fibrinopeptides have evolved rapidly with little constraint whereas histone IV has been evolving extremely slowly.

We attempted to estimate the variance of the evolutionary rate of a particular protein, in order to test the uniformity of the rate, and we found that the variance of amino acid substitution is larger than that expected if a simple Poisson process is assumed (Ohta and Kimura 1971). This analysis may be flawed because it is based on paleontological estimates of divergence time. But the conclusion has been confirmed
by Langley and Fitch (1974), who used more data and statistics less dependent on paleontological estimates. Kimura (1983) used data from mammalian orders that are believed to form a star phylogeny, i.e. simultaneous divergence of many mammals about 80 million years ago. For hemoglobins, myoglobin, cytochrome c and ribonuclease, the ratio, R, of the variance to the mean divergence turned out to be from 1.3 to 3.3, again confirming the previous result. Gillespie (1986) extended such analyses for more proteins; hemoglobins, cytochrome c, insulin, prolactin, ribonuclease, LHP, albumin, cytochrome oxidase, ATPase 6 and cytochrome b, and he found that R takes values between 0.2 and 34.1. The extreme values are thought to be due to small sample size (three species with star phylogenies were examined). Gillespie (1986) argues that this pattern fits the "episodic" process in which a burst of amino acid substitutions is followed by a static phase.

Since methods of DNA sequencing have become available, more and more data on DNA sequences are accumulating and molecular evolutionary studies have shifted from analysis of amino acid sequences to that of DNA sequences. Several remarkable features of DNA evolution have emerged. The majority of genomic DNA of higher organisms evolves more rapidly than protein coding regions, i.e. those DNA regions that apparently do not carry genetic information in their primary structure are evolving rapidly. In mammals, the rate of nucleotide substitution in these regions is roughly $5 \times 10^{-9}$ per site per year (e.g. see Kimura 1968; Nei 1987; Li and Graur 1991). The rate of synonymous substitutions in coding regions is slightly lower than this but is rather uniform among various genes, whereas the rate of amino acid replacement substitutions differs greatly from gene to gene. These values agree with the results of DNA hybridization studies (Britten 1986; Sibley and Ahlquist 1984).

An as yet unsettled problem is whether the rate of DNA evolution depends upon generation length. An examination of amino acid sequences revealed little effect of generation time, however DNA hybridization studies indicate that the longer the generation, the lower the DNA evolution rate (Laird et al. 1969; Kohn 1970; Britten 1986; Sibley and Ahlquist 1984). This is still a controversial problem and is discussed later.

Another topic of interest is the isochore concept, the differentiation of chromosomal regions of warm blooded vertebrates into GC- and AT-rich segments (Bernardi et al. 1985; Ikemura 1985). Such segments are called "isochores" and have an average size well above 200 kilobases (Bernardi et al. 1985). A noteworthy fact is that most housekeeping genes locate in GC-rich isochores, whereas many tissue-specific genes are in AT-rich ones. In addition, codon usage is different in these two groups of genes (Ikemura 1985). Thus, a codon bias in mammals appears to be largely determined by mutation pressure, but the bias in lower organisms with intermediate
GC content such as in *E. coli* is influenced by selection because of the efficient selection in very large populations (Sharp 1989). Such selection must be nearly neutral, i.e. at the border between neutrality and selection.

**Population Dynamics of Mutant Genes**

A basic requirement for understanding the mechanisms of nucleotide or amino acid substitutions in evolution is to distinguish mutations from evolutionary substitutions. Numerous mutations appear in Mendelian populations in every generation, but the majority will be lost within a few generations. Thus, those mutations that contribute to evolution are a very small minority of all mutations. It is also necessary to understand the process of frequency increase of mutants in the population in the course of their substitution.

For neutral mutants, the process has been theoretically analyzed using the diffusion equation method (Kimura and Ohta 1969; 1971). A neutral mutant, if it is ultimately fixed in the population, takes on the average $4N$ generations until this occurs.

$$t_i = 4N \quad (1)$$

where $N$ is the effective population size. If $N$ is large, the time is very long. The rate of molecular evolution is measured by averaging the number of substitutions over very long period of time. It may be expressed as follows,

$$k = \lim_{T \to \infty} \frac{n(T)}{T} \quad (2)$$

where $T$ is the period, and $n(T)$ is the number of mutant substitutions in this period. Obviously, for $k$ to be measured accurately, $T \gg N$.

Now consider a locus encoding a protein. Let the rate of occurrence of base substitutions in this DNA region be $v_g$ per generation, and let $u$ be the probability of fixation of a mutant. Then in a population of $N$ individuals, the total number of mutations appearing in the population is $2Nv_g$ per generation, and a fraction $u$ of them spread through the population, so the rate of substitution per generation becomes,

$$k_g = 2Nv_g u \quad (3)$$

Here $u$ generally depends upon the magnitude of natural selection. It should be remembered that, at the molecular level, the number of nucleotide sites of a locus is so large that the probability of having identical mutations more than once is almost nil. Also, the probability of back mutation is negligibly small.
Let us now examine how natural selection influences the rate. If most substitutions are caused by Darwinian natural selection, and the average selective advantage of such substitutions is $s$ with no dominance, the fixation probability is roughly twice the selective advantage (Fisher 1930; Haladane 1932), and we have,

$$k_g = 4Ns$$

Hence $k_g$ depends on the product of three parameters, $N$, $s$ and $v_g$. But when most substitutions are selectively neutral, the fixation probability is equal to the initial frequency, $1/2 N$ (Kimura 1962), and we have,

$$k_g = v_g$$

In other words, the evolutionary rate is simply equal to the mutation rate, and it is independent of population size (Kimura 1968).

The actual rate of molecular evolution seems to be roughly constant per year for each protein as reviewed in the previous section. In order to explain this fact by a selection model (Eq. 4), one has to assume that parameters like $N$ and $s$ have been nearly equal in various lineages. Such a situation could hardly hold in very different environments, particularly in the lines leading to living fossils and in those with rapid phenotypic evolution. On the other hand, if the majority of substitutions are selectively neutral as was first proposed by Kimura (1968), from formula (5), the observed pattern may be explained if the neutral mutation rate is constant per year. Here the possible problems are: (i) how the generation length affects the evolution rate, (ii) how fluctuation in the evolutionary rate is related to natural selection, and (iii) how selective constraints influence the rate. These problems are discussed later in relation to weak natural selection, i.e., the interaction of random genetic drift and selection.

**Selective Constraints and Evolutionary Rate**

The rate of evolution is different from protein to protein, and the difference reflects the degree of constraint as explained before. This constraint is directly connected to the function and structure of protein or RNA molecules, i.e., the more rigid their function and structure are, the lower is the rate. This point has been beautifully shown for fibrinopeptides, hemoglobin, cytochrome c and histone IV by Dickerson (1971).

Such evolutionary features may be explained by the neutral mutation random drift hypothesis as follows. As in Kimura (1968), $v_T$ designates the rate of occurrence of new mutations in terms of nucleotide substitutions. Let $f_0$ be the fraction of such mutations that are selectively neutral, and the remaining $(1-f_0)$ be the fraction of
mutations that have a deleterious effect. Then the rate of neutral evolution becomes,

\[ k = f_0 \nu_r \]  \tag{6}

because deleterious mutations do not contribute to evolution. Favorable mutations are assumed to be too rare to have any statistical influence. Various degrees of selective constraint may be taken into account by \( f_0 \), e.g. for pseudogenes without constraint, \( f_0 = 1 \), whereas for amino acid replacement substitutions of histone IV, \( f_0 = 0 \). Here the question is whether or not mutations may be simply divided into neutral and deleterious classes. This problem is discussed later.

**Meaning of Near Neutrality**

In the previous sections, I have repeatedly emphasized that the rate of molecular evolution is strongly dependent upon selective constraints of proteins or nucleic acids. Under the neutral mutation-random drift theory, it is assumed that a certain fraction of new mutations are free of constraint or are selectively neutral, while the rest have deleterious effects and are selectively eliminated. An important question is how the two classes are distinguished by natural selection. Let us now examine theoretically the behavior of the mutants belonging to the borderline class. The critical quantity is the fixation probability of mutant genes. I examine the simplest case of a semidominant gene with selective advantage, \( s \). Fixation probability in finite populations has been shown to become the function of the product, \( Ns \), as follows (Kimura 1962),

\[ u = \frac{1 - e^{-4Ns}}{1 - e^{-4N}} \]  \tag{7}

where \( p \) is the initial frequency in the population, and is assumed to be much less than unity. Figure 1 shows the fixation probability of mutants as a function of \( Ns \) relative to the completely neutral ones (\( Ns = 0 \)). As can be seen from the figure, the fixation probability is a continuous monotone function of \( Ns \). Thus, when discussing molecular evolution by nearly neutral mutants, one has to consider all mutants around \( Ns = 0 \).

Let us examine the borderline mutations in some detail. Figure 2 shows classification of new mutations. The upper part shows the simple neutral model, and the lower part, the nearly neutral model. In the figure, neutral means strictly neutral, deleterious means definitely deleterious, and nearly neutral means intermediate between neutral and selected. The behavior of mutations of the nearly neutral class is affected by both selection and drift.

As shown in Figure 1, the effectiveness of selection is determined by the product
Fig. 1. The fixation probability \( u \) of a mutant as a function of the product of population size and selection coefficient, \( Ns \), relative to the completely neutral case. \( p \) is the initial frequency and the line, \( u = p \), is the value for a neutral mutant.

**Simple neutral theory**

<table>
<thead>
<tr>
<th>deleterious</th>
<th>neutral</th>
<th>advantageous</th>
</tr>
</thead>
<tbody>
<tr>
<td>selected</td>
<td>nearly neutral</td>
<td>neutral</td>
</tr>
</tbody>
</table>

**Nearly neutral theory**

Fig. 2. A Schematic diagram shows the proportion of various classes of mutations. Deleterious mutants are definitely deleterious and neutral mutants are strictly neutral, while most selected mutants are deleterious, but the group also includes advantageous alleles. Nearly neutral mutants comprise a class intermediate between neutral and selected mutants.
of the effective population size and the selection coefficient, \(N_s\). Actual species have various population sizes from very small to very large, and therefore the effectiveness of selection will differ among species. In addition, physiological conditions may influence weak selection, e.g. a constraint on an enzyme function may differ between homeotherms and poikilothersms. The simple separation of new mutations into the deleterious and neutral classes will then not be satisfactory. There should be a substantial fraction of mutations that belong to the nearly neutral class.

Especially in view of the importance of negative selection caused by constraints, it is likely that many nearly neutral mutants are very slightly deleterious, i.e. on the left side of \(N_s = 0\) in Fig. 1. My hypothesis of slightly deleterious mutants is based on this view and regards the near neutrality as the limit when the selection coefficient approaches zero (Ohta 1972; 1973; 1974; 1976). In other words, I propose that a substantial fraction of mutant substitutions at the molecular level are caused by random fixation of very slightly deleterious mutations.

The model of very slightly deleterious mutation is related to the molecular clock in an important way. One problem with the molecular clock is that the observed substitution rate is constant per year rather than per generation. The generation-time effect is particularly evident when divergence is measured by DNA hybridization (Laird et al. 1969; Kohn 1970; Britten 1986). According to these studies, the rate of divergence of single-copy genomic DNA varies from \(10^{-9}\) to \(10^{-8}\) per base pair per year depending upon the taxon, and there appears to be a high negative correlation between the rate and generation length. Similar effects have been found when synonymous and other unimportant DNA divergences are examined (Li et al. 1987; Wu and Li 1985). Thus, it seems that the rate of evolution of the majority of genomic DNA depends upon generation length, whereas the divergence rate of amino acid sequences is relatively insensitive to it.

In the early 1970s, I tried to explain this seemingly contradictory observation by using population genetics theory. The fundamental idea is that most genomic DNA of higher organisms can freely accumulate base substitutions, i.e. most new mutations are selectively neutral. On the other hand, amino acid substitutions are more likely to be influenced by natural selection, i.e. many of them may be regarded as nearly neutral, or very slightly deleterious (Ohta 1972). Let us now examine how slightly deleterious mutations are related to the generation-time effect. The left side of \(N_s = 0\) in Fig. 1 shows that there is a negative correlation between the fixation probability and population size, provided that the selection coefficient is unchanged by population size. In other words, the chance of spreading by random drift is much higher in a small population than in a large population.

This prediction has an important bearing on the generation-time effect. In
NEARLY NEUTRAL THEORY OF MOLECULAR EVOLUTION

general, large organisms have a long generation length and small population size and vice versa. Thus there is a negative correlation between population size and generation length (Chao and Carr 1993). On the other hand, the mutation rate per year will be lower in organisms with longer generation times as it is ordinarily believed (Watson 1976, p. 254), and DNA divergence reflects such an effect. With selection this effect is expected to be partially cancelled by the negative correlation between the substitution rate and population size. Thus the amino acid substitution rate is expected to be relatively insensitive to generation length as compared to DNA divergence because amino acids are more likely to be subject to negative selection than are noncoding regions. In a more realistic model assuming normal distribution of mutants' effect, the similar relationship between the evolutionary rate and population size is predicted (Ohta and Tachida 1990; Tachida 1991).

Testing the Theory by DNA Sequence Analysis

I obtained nucleotide sequences from the genetic databases maintained at the National Institute of Genetics, which include GenBank, DDBJ (DNA Data Bank of Japan) and EMBL. The sequences used in the analysis are listed in Table 1. They have been chosen with the following conditions: the coding region is not small, and the protein function has not changed for a long time. Note that, if the function changes,

<table>
<thead>
<tr>
<th>Proteins</th>
<th>Primates</th>
<th>Artiodactyla</th>
<th>Rodentia</th>
</tr>
</thead>
<tbody>
<tr>
<td>albumin</td>
<td>V00494</td>
<td>M73993</td>
<td>J00698</td>
</tr>
<tr>
<td>lactate dehydrogenase A</td>
<td>X02152</td>
<td>D90143</td>
<td>M27554</td>
</tr>
<tr>
<td>fibrinogen γ</td>
<td>X02415</td>
<td>X15556</td>
<td>X05860</td>
</tr>
<tr>
<td>acetylcholin receptor α</td>
<td>Y00762</td>
<td>X02509</td>
<td>X03986</td>
</tr>
<tr>
<td>acetylcholin receptor β</td>
<td>X14830</td>
<td>X00962</td>
<td>M14537</td>
</tr>
<tr>
<td>growth hormone</td>
<td>J00148</td>
<td>M27325</td>
<td>K03232</td>
</tr>
<tr>
<td>growth hormone receptor</td>
<td>X05662</td>
<td>X54429</td>
<td>M33324</td>
</tr>
<tr>
<td>prolactin</td>
<td>V00566</td>
<td>J00022</td>
<td>X02892</td>
</tr>
<tr>
<td>insulin-like growth factor 1</td>
<td>M37484</td>
<td>X15726</td>
<td>X06107</td>
</tr>
<tr>
<td>insulin-like growth factor 2</td>
<td>M17862</td>
<td>X53553</td>
<td>M14951</td>
</tr>
<tr>
<td>igf binding protein 1</td>
<td>Y00856</td>
<td>X54979</td>
<td>M89791</td>
</tr>
<tr>
<td>igf binding protein 3</td>
<td>M35878</td>
<td>M76478</td>
<td>M33300</td>
</tr>
<tr>
<td>interleukin 1α</td>
<td>M15329</td>
<td>M37210</td>
<td>D00493</td>
</tr>
<tr>
<td>interleukin 1β</td>
<td>M15840</td>
<td>M37211</td>
<td>M15131</td>
</tr>
<tr>
<td>interleukin 2</td>
<td>V00564</td>
<td>M12791</td>
<td>K02797</td>
</tr>
<tr>
<td>interleukin 6</td>
<td>M29150</td>
<td>X57317</td>
<td>J03783</td>
</tr>
<tr>
<td>interleukin 7</td>
<td>J04156</td>
<td>X64540</td>
<td>X07962</td>
</tr>
</tbody>
</table>
the ordinary pattern of synonymous vs. nonsynonymous substitutions will be violated (OHTA 1991).

For the acquisition and analysis of the data, I used the ODEN Package created by INA (1992) at the National Institute of Genetics. The method of NEI and GOJOBORI (1986) is used, which is incorporated into the ODEN package, for estimating the numbers of synonymous and nonsynonymous substitutions. This method divides the nucleotide substitutions into synonymous and nonsynonymous categories, and then the number of multiple hits is estimated under the assumption of random mutability among the four kinds of bases. Because this assumption is often not satisfied, the method is not completely satisfactory. Recently, LI (1993) and PAMLO and BIANCH (1993) invented a better method. In this method, the problem of bias in transitional vs. transversional substitutions is overcome by taking the weighted average of these changes, at twofold and fourfold degenerate sites, for estimating the number of synonymous substitutions.

INA (1993) has invented another more efficient method for estimating the numbers of synonymous and nonsynonymous substitutions. His method brings KIMURA's two-parameter model (1980) into the NEI-GOJOBORI method (1986). Through extensive simulations, INA has shown that his method gives even better estimates than LI's, since the transitional substitution rate often considerably exceeds random expectation.

From the estimated divergence of pairwise sequence comparisons among primates, artiodactyla and rodentia, the branch length of the star phylogeny was estimated as in KIMURA (1987). Let $d_{pa}$, $d_{pr}$, and $d_{ar}$ be the estimated divergences between primates and artiodactyla, between primates and rodentia, and between artiodactyla and rodentia respectively. Also let $d_p$, $d_a$ and $d_r$ be the branch length of primates, artiodactyla and rodentia lineages of the star phylogeny. Then, we have,

$$d_p = (d_{pa} + d_{pr} - d_{ar}) / 2$$
$$d_a = (d_{pa} + d_{ar} - d_{pr}) / 2$$
$$d_r = (d_{pr} + d_{ar} - d_{pa}) / 2$$

A total of 17 gene loci have been analysed, and phylogenies of these 17 loci are shown in Figure 3. The estimated total numbers of nonsynonymous and synonymous substitutions are shown beside each branch. The Roman figures are the numbers estimated by NEI and GOJOBORI (1986), the italic ones by LI (1993), and the gothic ones by INA (1993). The estimated total numbers do not differ much under the three methods. However, the estimated numbers of synonymous and nonsynonymous sites considerably differ by the three methods: the number of synonymous sites is larger, and that of the nonsynonymous sites is smaller, by LI's method as well as by INA's method that by NEI-GOJOBORI method. As the result, the number of synonymous substitutions
Fig. 3. Star phylogenies of 17 genes. The number beside each branch is the estimated total number of substitutions; the Roman figures, the value estimated by the method of NEI and Gojobori, the italic figures by that of Li, and the gothic ones by that of INA.

*per site* is considerably smaller with the Li's and the INA's methods than with the NEI-Gojobori method.

Li's previous method (Li et al. 1985) has also been tried, but the results are not given here because the estimated numbers are very close to the values obtained by NEI and Gojobori. According to INA (1993), the new methods of Li and of INA are more efficient than that of NEI-Gojobori, as they give very close values to the true divergences in INA's simulations.

From Fig. 3, it is also clear that the rodent line is the most divergent among the three lineages for synonymous substitutions, but not necessarily so for nonsynonymous substitutions. This is in accord with previous observations on other loci (Li et al. 1987; Gillespie 1991).

By using the result of Li et al. (1987), Gillespie (1991) estimated the lineage effect by a weighting factor. This factor is the characteristic divergence of each lineage and the average is constrained to equal one. From the branch lengths of the star phylogenies in Figure 3, I obtained weight factors for synonymous and nonsynonymous substitutions. The comparison of Gillespie's results and the present results is given in Table 2. The difference between the previous and the present estimates is not large, but the pattern of a large synonymous weight factor for the rodent lineage and a small
Table 2. Weight factors and the ratio of the number of nonsynonymous substitutions to that of synonymous substitutions per site.

<table>
<thead>
<tr>
<th>Weight factor</th>
<th>Rodentia</th>
<th>Artiodactyla</th>
<th>Primates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synonymous</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nei-Gojobori</td>
<td>1.694</td>
<td>0.757</td>
<td>0.549</td>
</tr>
<tr>
<td>Li 93</td>
<td>1.671</td>
<td>0.802</td>
<td>0.528</td>
</tr>
<tr>
<td>Ina 1</td>
<td>1.649</td>
<td>0.784</td>
<td>0.566</td>
</tr>
<tr>
<td>Gillespie (data by Li)</td>
<td>1.611</td>
<td>0.762</td>
<td>0.627</td>
</tr>
<tr>
<td>Nonsynonymous</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nei-Gojobori</td>
<td>1.313</td>
<td>0.905</td>
<td>0.782</td>
</tr>
<tr>
<td>Li 93</td>
<td>1.305</td>
<td>0.881</td>
<td>0.814</td>
</tr>
<tr>
<td>Ina 1</td>
<td>1.330</td>
<td>0.883</td>
<td>0.787</td>
</tr>
<tr>
<td>Gillespie (data by Li)</td>
<td>1.279</td>
<td>0.885</td>
<td>0.830</td>
</tr>
<tr>
<td>Ratio</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nei-Gojobori</td>
<td>0.162</td>
<td>0.251</td>
<td>0.300</td>
</tr>
<tr>
<td>Li 93</td>
<td>0.209</td>
<td>0.293</td>
<td>0.411</td>
</tr>
<tr>
<td>Ina 1</td>
<td>0.253</td>
<td>0.352</td>
<td>0.436</td>
</tr>
<tr>
<td>Gillespie (data by Li)</td>
<td>0.166</td>
<td>0.242</td>
<td>0.279</td>
</tr>
</tbody>
</table>

The pattern found for synonymous substitutions represents generation-time effects. One can note that generation-time effects are smaller for nonsynonymous substitutions than for synonymous substitutions. In fact, the ratio of the weight factor for rodent to that for primate is about 3 for synonymous substitutions, but only 1.6 for nonsynonymous substitutions. This result is in accord with the nearly neutral theory of molecular evolution (Ohta 1972; 1973; 1992). In other words, if one assumes that most synonymous substitutions belong to the neutral class, whereas most nonsynonymous substitutions belong to the nearly neutral class, the present result may be easily explained. Note that a negative correlation is expected between evolutionary rate and population size for nearly neutral mutations, and that rodent species are thought to have larger population sizes than primate species (Ohta 1972; 1992).

In Table 2, the ratio of the numbers of nonsynonymous substitutions to those of synonymous substitutions per site is also given. Note that this ratio is the proportion of acceptable amino acid substitutions. The estimated value of the proportion varies among lineages, and also by statistical methods. The highest estimate is obtained by the method of INA, which gives the most satisfactory result in simulation studies (INA 1993).

The present result shows that the proportion of acceptable amino acid substitu-
tions in the primate lineage is about twice as large as that in the rodent lineage. The proportion in the artiodactyl lineage is between the two values but close to that of the primate lineage. One has to note here that this result holds for those genes whose function has been fixed for a long time and does not apply to duplicated genes and others whose function has been modified in the evolutionary course of the lineages studied (Ohta 1991). It is likely that slightly advantageous amino acid substitutions to increase functional diversity among duplicated genes contribute to the evolutionary pattern of incipient gene families.

Summary

Various studies on population genetics and molecular evolution are reviewed with special reference to the nearly neutral theory, which proposes that a substantial fraction of new mutations at the coding region are nearly neutral. Based on this theory, it was predicted 20 years ago that the rate of DNA evolution depends on generation length more strongly than does that of protein evolution. The prediction is tested through analysis of sequence data of 17 mammalian genes by estimating divergence among genes separately for synonymous substitutions and nonsynonymous substitutions. Star phylogenies composed of rodentia, artiodactyla and primates are examined. The generation-time effect is found to be more conspicuous for synonymous substitutions than for nonsynonymous substitutions. The proportion of accepted amino acid substitutions is estimated to be about twice as large in the primate lineage than in the rodent lineage. This result supports the nearly neutral theory.

References

分子進化のほぼ中立理論とDNAデータ解析によるその検証

太田朋子
国立遺伝学研究所
〒411 静岡県三島市谷田1111

生物の表現型の進化はダーウィンの自然淘汰説によって説明されるが、これは定性的であって定量的ではない。集団遺伝学は半世紀以上前に進化機構を定量的に理解しようという意図のもとに始まった。進化は生物の種のすべての個体の変化であるという認識から、集団（種）の遺伝子頻度がどのように変化するかが解析された。しかし表現型に基づくかぎり進化を遺伝子頻度の変化と結びつけることはきわめて困難であった。分子生物学の進歩により遺伝情報がDNAの線状構造にあり、これが世代から世代へと伝えられることがわかった。そしてDNAや蛋白質のアミノ酸配列を各種生物の間で比較することができるようになった。このような進歩によって伝統的なネオダーウィニズムのいくつかの概念は見直しを迫られることとなった。見直しの第一歩は木村資生博士が1968年に提唱した分子進化の中立説である。

蛋白質の進化速度は、機能的に重要なものを進化が遅い。中立説によれば、重要な蛋白質では、突然変異の中で有害効果をもつものの割合が高いが、重要でない蛋白質ではこの割合が低い。すなわち中立突然変異の割合の高低によって進化速度の違いが生じるとする。そこで問題があるが、有害突然変異と中立突然変異の中間すなわち淘汰をうけるかどうかいかの境界はどうなっているかという点である。私は20年程前のどのような境界をなからほど中立突然変異による進化について検討した。ここでは遺伝的浮動の働き方が問題になるわけて最も重要なパラメターは集団の有効な大きさである。集団遺伝学の理論的研究から次のようなことがわかった。完全中立な突然変異では進化速度は集団の大きさとは無関係である。古典的な自然淘汰説によると進化速度と集団の大きさとは正の相関がある。これに対し中立的な突然変異では進化速度と集団の大きさとは負の相関がある。この理論は分子進化時計が世代あたりでなく年あたり成立するように見えるのはなぜかという問題と関連している。もしアミノ酸置換の多くは中立理論に従うとすれば、アミノ酸置換は、ゲノム全体からみたDNAの進化に比べ世代依存性が少ないと予測できるからである。

現在ではDNAの塩基配列を各種生物で比較したり、あるいは重複した遺伝子の間で比較したりしてこの予測を検討することができる。哺乳類の基本的な機能を持つ17の遺伝子について進化速度を計算してきた。同義置換および非同義置換において計算すると、同義置換についてはネズミなど世代の短い動物で速いといった世代効果が著しいのに対し、非同義置換では世代効果がはっきりしないことがわかった。ネズミは体が小さく齀長類に比
べ集団が多いと考えられる。したがって集団が多いとアミノ酸置換が同義置換に比べ少ないということになる。先の理論的予測のもとにアミノ酸置換の多くがほぼ中立であるとすれば理解しやすい。この世代効果は主として機能を保持するような淘汰が働いた場合に見られる。重複遺伝子その他新しい機能を獲得したり、機能が修正されたりする場合は、より積極的な淘汰が働き世代効果は異なってくる。
GENETIC VARIATION AND GENE REGULATION
AT THE WX LOCUS OF RICE

Hiro-Yuki Hirano
National Institute of Genetics
1111, Yata, Mishima-shi, Shizuoka-ken 411

Introduction

The \( \text{wx} \) locus controls amylose synthesis in the endosperm and pollen. Since mutations at this locus are easily identified as altered phenotypes of such organs, many genetic studies on this locus have been carried out. Especially, transposable elements were discovered at the loci including the \( \text{wx} \) locus of maize for the first time among all organisms (for review; Fedoroff 1983). At the molecular level, the first plant transposon (\( \text{Ac} \) element) was cloned from the \( \text{wx} \) locus (Shure et al. 1983). The fine structure, which was the map of mutation sites within a gene, was constructed genetically by pollen analysis of this locus (Nelson 1968). Thus, the \( \text{wx} \) locus has contributed to the brilliant studies in the history of plant genetics.

The \( \text{wx} \) locus encodes ADP glucose starch glycosyl transferase, which catalyzes the synthesis of amylose in the endosperm (Nelson and Rines 1962). The mutant of this locus is unable to produce amylose and starch in the endosperm is almost composed of amylopectin. In rice, amylose content (relative to total starch) greatly affects the quality of rice when cooked. For example, rice that contains amylose about 14–18% is moderately glutinous and is suitable for "boiled rice", which is a staple diet for Japanese people. On the other hand, rice that has amylose more than 25% is not so sticky as its grains cannot aggregate to form a clump.

Thus, the \( \text{wx} \) locus has an advantage to investigate the regulation of gene expression because the genetic approach is available. In addition, the elucidation of gene regulation of the \( \text{wx} \) locus may provide some important findings to be applied for the molecular breeding to improve rice quality. We have cloned the DNA from the rice \( \text{wx} \) locus and prepared the antisera against its gene product (Wx protein). Using these molecular markers, we have been studying the gene expression at the \( \text{wx} \) locus of rice. In this symposium, we would like to report our present work on the Wx gene expression and to discuss its regulation and use for rice breeding.
Results and Discussion

Nucleotide sequence and gene structure of the Wx gene

We cloned the Wx gene from rice (O. sativa Japonica; Taichung 65 (T65)) and determined the entire sequence of the coding region (HIRANO and SANO 1991a). The rice Wx gene was divided into a number of exons with introns as was the maize Wx gene. The smallest exon was 64 bp in length and the largest intron was 357 bp in length in the rice gene. The deduced amino acid sequence indicated a putative transit peptide of 77 amino acids and a mature protein region of 532 amino acids. The putative transit peptide may play an important role for transporting Wx protein into the amyloplast (KLÖSGEN 1991).

The nucleotide sequences of Wx gene from maize (KLÖSGEN 1986) and barley (ROHDE 1988) has been determined. The exon–intron structures of the Wx gene from rice, maize and barley are quite similar each other, although the barley Wx gene lacked two introns (Fig. 1) (see later section). The amino acid sequences of the mature Wx proteins including that of wheat (AINSWORTH 1993) showed high homologies (Table 1).

The waxy phenotype and its genetic description is restricted to plants in

![Gene structures for the wx locus and GBSS. Boxes with shading indicate exons and open boxes indicate introns.](image)

**Table 1. Homologies in deduced amino acid sequences (%).**

<table>
<thead>
<tr>
<th></th>
<th>Maize</th>
<th>Barley</th>
<th>Wheat</th>
<th>Potato</th>
<th>Pea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rice (wx&lt;sup&gt;+&lt;/sup&gt;)</td>
<td>88</td>
<td>87</td>
<td>85</td>
<td>69</td>
<td>67</td>
</tr>
<tr>
<td>Maize (wx&lt;sup&gt;+&lt;/sup&gt;)</td>
<td>86</td>
<td>85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barley (wx&lt;sup&gt;+&lt;/sup&gt;)</td>
<td></td>
<td></td>
<td>95</td>
<td>69</td>
<td>66</td>
</tr>
<tr>
<td>Wheat (wx&lt;sup&gt;+&lt;/sup&gt;)</td>
<td></td>
<td></td>
<td></td>
<td>68</td>
<td>64</td>
</tr>
<tr>
<td>Potato (GBSS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>75</td>
</tr>
</tbody>
</table>
Gramineae, except for grain amaranths (KONISHI et al. 1985). However, granule-bound starch synthase (GBSS) in potato (VAN DER LEJ et al. 1991) and pea (DRY et al. 1992) have biochemical properties quite similar to those of Wx proteins of monocotyledonous plants, that is, GBSS has enzymatic activity for ADP glucose starch glycosyl transferase and localizes in the starch granule. The putative amino acid sequences of GBSS from potato and pea showed sequence similarity to those of Wx protein (Table 1). The exon–intron structure of the gene for potato GBSS was the same as that of the maize Wx gene. Therefore, the Wx gene in Gramineae and the genes for GBSS in potato and pea may belong in the same gene family and may have been derived from a common ancestral gene during evolution.

Tissue Specific Gene Expression

We examined the tissue specific expression of the rice Wx gene. Northern analysis showed that a 2.4 kb transcript from the Wx locus in rice was present only in the seed, and not in the leaf, stem, or whole seedling without seeds. Thus, the Wx gene in rice was expressed in a tissue–specific manner. The immuno-blot analysis using the antibody against Wx protein showed that the Wx gene was expressed not only in the endosperm but also in the pollen, although the amount of the Wx protein in the pollen was about 2% of that in the endosperm. Thus, the gene was expressed only in the endosperm and pollen, and the regulation was at the transcriptional level (HIRANO and SANO 1991a). Of great interest to us is the molecular mechanism of gene regulation in two distinctly different, differentiated tissues, the function, developmental origin and ploidy of which are different.

In order to examine the promoter function for tissue specificity, we introduced the chimeric DNA consisting of the 5’ upstream region of the rice Wx gene and bacterial β-glucuronidase (GUS) gene as a reporter into cells of rice and regenerated whole plants from them (HIRANO et al. manuscript in preparation). The functions of the 5’ upstream region were monitored by histochemical analysis of GUS activity in various tissues (organs) of transgenic plants. In the seed of transgenic rice, GUS activity was located in the endosperm and no activity was detected in the embryo (Fig. 2). GUS activity was also detected in the pollen but was not in the leaves, stems and roots. These results indicated that the 5’ upstream region introduced was sufficient to direct the tissue–specific expression in the endosperm and pollen.

On the other hand, in the case of transgenic petunia, the same chimeric gene was expressed in the pollen but was not properly expressed in the endosperm. These results suggest that the cis–acting elements that drive the gene in the pollen are common to both monocotyledonous and dicotyledonous plants but those for endosperm–specific expression in rice, did not function in dicotyledonous plants.
Genetic variation and quantitative gene regulation at the \textit{wx} locus

We examined the amount of Wx protein in the endosperm from various species with the AA genome in \textit{Oryza} genus by immunoblot analysis. Results indicated that rice were classified into two groups by the amount of Wx protein; \textit{O. sativa} Japonica (Fig. 3, lane 2-5) and the other rice species including \textit{O. sativa} Indica (Fig. 3, lane 6-10). The amounts of Wx protein of the latter group were about 5-10 fold than those of \textit{O. sativa} Japonica. In order to know whether this quantitative difference is due to the function of the \textit{wx} locus or to the genetic background, we analyzed the near-isogenic line that have the \textit{wx} locus from \textit{O. glaberrima} or \textit{O. rufipogon} on the genetic background of \textit{O. sativa} Japonica (T65). The result indicated that these near-isogenic rice produced the same amount of Wx protein as their donor species. Therefore, we concluded that the \textit{wx} locus regulated the quantitative level of the Wx protein synthesis in each species (or strain). Judging from the quantitative capacity for the synthesis of Wx protein, we proposed the idea that the rice \textit{wx} locus had two wild-type alleles, \textit{Wxa} and \textit{Wxb}; \textit{Wxa}, which is distributed in various species including \textit{O. sativa} Indica and wild rice, produced Wx protein 5–10 fold than \textit{Wxb} in \textit{O. sativa} Japonica did (SANO 1984).

Northern blot analysis showed that the amount of transcript from \textit{Wxa} (an alleles of patpak (Indica) on the genetic background of T65 (Japonica)) was about 8-fold than that from \textit{Wxb} in T65 (HIRANO and SANO 1989). Thus, quantitative control of Wx
Fig. 3. Immuno–blot analysis of Wx proteins from various rice strains (species). Total protein extract from mature seeds were analyzed by the immuno–blot analysis using anti-serum raised against Wx protein. *O. sativa* Japonica: (1) 563 (the wx mutant), (2) T65, (3) Norin 8, (4) 535, (5) 552; *O. sativa* Indica: (6) 868, (7) 108, (8) 414; *O. rufipogon* (9) W593; *O. glaberrima* (10) W025.

protein was at the transcriptional level. The molecular basis for distinguishing the \( Wx^a \) and \( Wx^b \) may be due to the strength of the promoter of the Wx gene or enhancer-like element(s) near this gene.

Since *O. sativa* Japonica, cultivated rice, has been evolved from the ancestral wild rice *O. rufipogon* (Chang 1984; Oka 1988), \( Wx^b \) is presumably derived from \( Wx^a \). It is of great interest what kind of mutations had altered \( Wx^a \) to \( Wx^b \) during the domestication process of *O. sativa* Japonica.

Since Wx protein catalyzes amylase synthesis, its quantitative level should strongly affect amylase content in mature seeds. In fact, amylase contents in rice with \( Wx^b \) were 14–16%, while those with \( Wx^a \) were 22–26% (Sano 1984). The correlation of the amount of Wx protein and amylase content will be discussed later in detail.

**Gene dosage effect**

The Wx gene expression was strongly dependent on gene dosage. We estimated
the amount of the Wx protein from progenies of reciprocal cross of T65 (wild type) and T65 \( wx \) (a mutant of \( wx \) locus). The results indicated that the amount of Wx protein were proportional to the number of wild-type alleles. In the case of cross between \( Wx^a \) and \( Wx^b \), the amounts of Wx protein from plants with genotype of \( bbb, abb, aab \) and \( aaa \) were approximately \( 1 : 3 : 6 : 9 \). This quantitative regulation may be at the transcriptional level, because the same results were obtained in Northern blot analysis. These results strongly suggested that the two wild-type alleles independently produce the transcripts according to their quantitative capacity and the two alleles did not interact each other even if they were present together with in a nucleus.

The function of \( du \) loci

The low amylose mutants, which were first found by Okuno et al. (OKUNO 1983) and designated \( dull (du) \), contain lower amounts of amylose in the endosperm than wild-type strains. We have screened and obtained mutants similar to \( du \) (SANO 1985) and analyzed the function of these \( du \) loci with reference to the gene expression of \( wx \) locus (HIRANO et al. manuscript in preparation).

The loci of six mutants obtained were independent not only of the \( wx \) locus but also of each other and they had lower amylose content (5%-9%) than that of the original wild strain (Norin 8; 14%). Immuno-blot analysis with antibody for Wx protein indicated that the amounts of Wx protein were reduced in all \( du \) mutants although the extent of reduction is variable (Fig. 4). The relationships of relative amount of Wx protein and amylose content in \( du \) mutants are shown in Fig. 6 (closed circulars). Northern blot analysis showed that this reduction of Wx protein level was controlled at the transcriptional level. These results strongly suggests that the \( du \) loci are responsible for positive regulation of the Wx gene expression. Therefore, the mutation at the \( du \) loci may reduce the production of the Wx protein, which catalyzes the amylose synthesis, lowering amylose content in the endosperm.

If the \( du \) loci encode trans-acting factors that activate the transcription of the Wx gene, it is expected that they affect expression of some other genes. Based on this hypothesis, we examined total peptides translated in vitro from poly A+RNA isolated from developing seeds by two dimensional gel electrophoresis. We found that the quantitative levels of Wx protein and a few additional peptides decreased in \( du \) mutants. On the contrary, the levels of a few other peptides increased in the same \( du \) mutants. These results suggest that \( du \) loci may regulate the expression of some genes besides the Wx gene and may have both function for activating and repressing the genes. The function, which may be positive regulation or negative, may depend on the genes affected. It is of a great interest what kind of genes are controlled by the \( du \) loci and what mechanisms underlie this regulation.
Fig. 4. The quantitative levels of Wx protein from a wild strain (1) and du mutants (4–5). Total protein extract from mature seeds were analyzed by the immuno-blot analysis. (1) Norin 8, (2) 2035, (3) 76–3, (4) 74–5, (5) 75–2.

Cool temperature response

When rice plants are exposed to cool temperatures (less than 20°C) during seed development, the amylose contents in the endosperm increase causing a lower quality of rice when cooked. We examined this response to cool temperatures with respect to gene expression at the wx locus (SANO et al. 1985; HIRANO and SANO 1991b).

Rice plants were grown during seed maturation at two different temperatures, 28°C (normal) and 18°C (cool), after anthesis. The amounts of the gene product (Wx protein) of the wx locus were estimated by Immuno-blot analysis. The result indicated that not only amylose content but also the amount of Wx protein increased in seeds matured at 18°C compared with those at 28°C. This suggests that the increase of amylose content is brought about by an elevated level of Wx protein, which catalyzes amylose synthesis.

Plants were grown at 18°C for the different lengths of time, during the period in which Wx protein accumulated linearly, that is, for 1, 2, 4, 8 or 12 days from the 7th day after anthesis, and then grown at normal temperatures (in the field) until seed
maturation. The longer plants were exposed to a temperature of 18°C, the higher levels of amyllose and Wx protein accumulated in mature seeds. The extent of increase in the amyllose content was positively correlated with that in the Wx protein level (Fig. 6; closed triangles). This result strongly supports the idea that the amyllose contents increase at cool temperatures because of elevated levels of Wx protein, the enzyme for amyllose synthesis.

Northern and slot–blot analysis indicated that the steady state level of the Wx transcript increased about 2.5-fold by exposing the plants to a temperature of 18°C (Fig. 5, slot a and b) and it decreased to the initial level after the plants were grown at 28°C for 2 days (Fig. 5 slot b–e). These results suggest that the response to a cool temperature is regulated at the transcriptional level of the Wx gene expression and this regulation is reversible.

The regulation in response to a slight difference of temperatures may be very

---

**Fig. 5.** Response of the Wx gene expression to cool temperatures. (A) Slot–blot hybridization. (B) Schematic representation of the experimental procedures. Rice plants with developing seeds were exposed to two temperatures (18 and 28°C) as indicated. The seeds were harvested at the time indicated (a–g) and quickly frozen in liquid nitrogen. Total RNAs were isolated from them and the steady state levels of the Wx transcript were estimated by slot–blot hybridization using the Wx cDNA as a probe.
important for the growth and survival of plants. The study on the \( Wx \) gene in response to a cool temperature may provide us one of the best clues for investigating this regulation.

**The amount of \( Wx \) protein and amylose content**

From the results of studies described above, the correlations between the relative amount of \( Wx \) protein and amylose content are summarized in Fig. 6. Amylose content was proportional to the relative amount of \( Wx \) protein in the region A. As the \( Wx \) protein is higher, the correlation curve was saturated. The similar results were
obtained by the analysis of the \textit{du} mutants and natural rice strains (Sano 1985). We interpreted these results as follows: (i) the amylose might be synthesized in proportion to the amount of \textit{Wx} protein, which catalyzes the amylose synthesis (ii) since the amylose occupied quite large portion in the endosperm, its synthesis might be restricted with physical and spatial constraints in spite of a lot of \textit{Wx} protein in the region B.

\textbf{Perspective for molecular breeding}

The techniques for manipulating DNA and making transgenic plants have enabled us to improve various crops, for example, resistant to viruses (Beachy 1990). We may also alter the amylose content in rice seeds by controlling the expression level of the \textit{Wx} gene in transgenic rice. There are two ways for this control. The first way is introducing the chimeric DNA consisting of the promoter region of the rice \textit{Wx} gene and cDNA for \textit{Wx} protein into the \textit{wx} mutant. The other way is introducing the chimeric DNA, which is the same combination as above but produce anti-sense RNA, into wild-type rice. We expect that the former method is more hopeful, because it can control delicately the levels of the \textit{Wx} gene expression. For the delicate control, it is quite important to characterize the mechanisms responsible for the quantitative regulation of gene expression. Putative \textit{cis}-elements that defined alleles of \textit{Wxa} and \textit{Wxb} may be an essential clue to elucidate such mechanisms.

It is also important to find the \textit{cis}-element responsible for the response to cool temperatures. If such elements are identified, we can produce the transgenic rice, whose seeds synthesize amylose at an appropriate level even though they develope at cool temperatures, by mutagenizing or deleting these elements.

In general, expression level of genes introduced are different each other among transformants. This phenomenon, so called a position effect, is one of the major problems not only for basic studies of gene regulation but also for molecular breeding for improving crops. We expect that this problem may be solved in future by the techniques for ‘gene replacement’ or ‘locus-directed integration’ of genes to be introduced.

\textbf{Intron loss during evolution of Gramineae}

As described earlier, two introns (X and Y) are absent in the \textit{Wx} gene from barley (Röhde 1987), compared with the \textit{Wx} gene in rice (Hirano and Sano 1991) and maize (Klösgen 1986). While potato (van der Leij et al. 1991), which is a dicotyledonous plant, is distant from \textit{Gramineae}, the GBSS gene of potato has both of the two introns at the same position as the \textit{Wx} gene of rice or maize. This suggests that the two introns had been present in an ancestral gene of the \textit{Wx} gene and GBSS gene and they
were lost from the Wx gene in barley during evolution (HIRANO and SANO 1992).

To examine when the two introns were lost during evolution of barley, we began to examine their distribution in various Gramineae species. The presence or absence of these intron was analyzed by PCR methods using primers in their flanking exons and confirmed by Southern analysis of the PCR products. The results from preliminary experiments showed that the intron X was lack in all the seven species examined in the tribe Triticeae, while it was present four species in the other tribes, Agrostieae and Festucae, in the same subfamily, Pooideae. This suggests that the intron X had been lost during the divergence of ancestral species in the Pooideae subfamily. Since similar approaches succeeded in estimating the time of retrotransposition of a plant SINE (HIRANO et al. 1994), we expected that the further analysis may reveal when these introns were lost from the wx locus during evolution of Gramineae.

Two arguments, the 'intron-early' and the 'intron-late' views, are proposed for origin of introns and controversial among scientists. The former view argues that introns are as old as the genes them selves and that the different numbers of introns of the same gene from species to species, are the result of the loss of intron (GILBERT et al. 1986; TITTIGER 1993). On the contrary, the latter view is that introns were mobile element and were inserted in pre-existing genes during evolutoin (PALMER 1991). Our results presented here support the former 'intron early' view, since it is very difficult to explain that introns were independently integrated into the same position in the genes of various species.

Summary

The wx locus in rice has some advantages for genetic research and many knowlages and genetic materials have accumulated. In addition, this locus controls amylose synthesis, whose contents (relative to total starch) are quite important for a rice quality when cooked. We have cloned the gene at the wx locus and have been studying its gene regulation. The main results obtained so far are as follows: (i) The wx locus had two wild type alleles, Wx^a and Wx^b, which are distinguished by the quantitative capacities to produce Wx protein. This quantitative differences were regulated at the transcriptional level. (ii) The du loci, whose mutation reduced the accumulation of amylose in the endosperm, may have the function that positively regulate the gene expression at the wx locus. (iii) The levels of the Wx gene expression were activated at a cool temperature (18°C), causing high amylose content in mature seeds and lowering rice quality. The results obtained here indicated that the amylose level accumulated in mature seeds strongly depended on the amount of Wx protein, suggesting that we could regulate the amylose contents at will by controlling the expression
levels of the Wx gene in developing seeds of transgenic rice.

Acknowledgements

I thank Dr. SANO in our laboratory for valuable advises and encouragements and for supplying genetnic materials. We thank following colleagues for technical helps and valuable suggestions: Drs. N. KAWAKAMI, K. NODA (Yokohama City Univ.), Y. KOMEDA (Univ. Tokyo), T. MATSUMURA, N. TABATYASHI, H. SARUYAMA, M. TANIDA (Hokkaido Green-Bio Inst.), H. KIKUCHI (Hokkaido Cent. Agr. Exp. Stn.), H. MORISHIMA and M. EGUCHI (Natl. Inst. Genet.). This work was supported in part by grant-in-aid for scientific research from the Japanese Ministry of Education, Science and Culture.

References


インネ wx 座の遺伝的変異と発現調節

平野博之

国立遺伝学研究所
〒411 静岡市清水区1111

wx 遺伝子座は遺伝学的な研究に有益であり、これに関する多くの知見や植物材料が蓄積している。この遺伝子座はインネの種子胚乳中のアミロースの合成を支配している。アミロース含量（全デンプンに対する割合）は米の品質に大きな影響を与えるため、この遺伝子座の発現制御は応用的にも重要である。我々はインネの wx 座の遺伝子をクローニング化し、その構造を解析するとともに、この cDNA や Wx タンパク質に対する抗血清を用いて、この遺伝子の発現様式やその調節機能を研究してきた。今まで得られた結果は次のとおりである。(i) wx 座には、強調の発現量から判断して二つの対立遺伝子 (Wxa と Wxb) があり、これらの違いは転写のレベルで制御されている。 (ii) du 遺伝子座に起こっ変異は低アミロースを引き起こすことが知られている。du 座は wx 座の発現を正に制御し、この遺伝子座の欠損によりアミロースを合成する酵素である Wx タンパク質が減少しその結果アミロース含量が低下する。 (iii) wx 座の発現は、18°C という弱低温で活性化され、Wx タンパク質の量が増加する。その結果、完熟種子中のアミロース含量が增大し、米の品質低下をもたらす。さらに、これらの結果を総合すると、完熟種子中のアミロース含量は Wx タンパク質の量によって調節されていると結論される。これは、トランスジェニックインネにおいて、導入された Wx 遺伝子の発現を調節することにより、その種子のアミロース含量を任意に制御することが可能であることを示唆している。

＜質疑応答＞

河合（三和生薬）：Tissue specificity についてのお考えをお聞かせ下さい。

平野：それぞれの組織で発現しろというような命令を制御する遺伝子があり、例えば胚乳では胚乳で発現しろという遺伝子群、それに応答する cis のエレメント（プロモーター領域）がそれに応答して発現するのだろうと考えています。

菅原（北海道グリーンバイオ）：モチについては Wxa と Wxb に対応するものがありますか。

平野：Oryza glaberrima のモチはあり、Wx タンパク質が出来てないということでは全く同じであります。しかし、glaberrima のモチはジャポニカのモチとはデンプンの質がかなり違うが、これは genetic background が違うためと考えています。

菅原：アミロース含量は登熟気温によって変化することですが、du 遺伝子を持つ方が温度反応性が高いと思うが、これについてはどうですか。
平野：全ての dull で温度反応するようだ。温度反応しない dull があればおもしろい。
中島（生物研）：du 遺伝子の発現の温度反応は調べていますか。
平野：まだやっていません。穂を切っても調査できるので、これから研究します。
MOLECULAR EVOLUTION OF CONIFERS

Yoshihiko TSUMURA and Kihachiro OHBA

1 Bio-resources Technology Division, Forestry and Forest Products Research Institute, Kukizaki, Ibaraki, 305 Japan
2 Institute of Agriculture and Forestry, University Tsukuba, Ibaraki, 305 Japan

Introduction

Genetic feature and phylogeny in conifers have been studied using mainly isozyme analysis since 1970’s. The major facts which have been revealed by isozyme study are, 1) the average of outcrossing rate is more than 90% (ADAMS and BIRKES 1991), 2) genetic variation is higher ($He = 0.173$) than those of the other plant species ($He = 0.149$), and 3) genetic differentiation between population within species is quite low ($G_{ST} = 0.068$) (HAMRICK and GOET 1989). Japanese conifer species also have similar values in their variation and diversity (TSUMURA and OHBA 1992, 1993). Recently, as molecular biology technique has developed, the genetic features of conifer in DNA level have also been revealed since the middle of 1980’s. Chloroplast DNA (cpDNA) has been studied extensively, which has a genome structure (STRAUSS et al. 1988, LIDHOLM et al. 1988, LIDHOLM and GUSTAFSSON 1991, TSUZUKI et al. 1992, TSUMURA et al. 1993, WHITE et al. 1993), hereditary traits (NEALE et al. 1986, SZMIDT et al. 1987, NEALE and SEDEROFF 1989, NEALE et al. 1989), and intra- and inter-specific variations (WANGER et al. 1987, WHITE 1990, SUTTON et al. 1990, ALI et al. 1991, WAGNER et al. 1992). For mitochondrial DNA (mtDNA), its inheritance mode has been investigated in some species (NEALE et al. 1989, 1991). Analyses of Restriction Fragment Length Polymorphism (RFLP) and Random Amplified Polymorphic DNA (RAPD) for the linkage mapping and the clone identification also have been stratified (DEVEY et al. 1991). Molecular phylogenetic relationships of conifer species also have been studied; namely Pinus species using RFLP of cpDNA, mtDNA, and other (STRAUSS and DOERKSEN 1990), Pinus species using ribosomal DNA RFLP (GOVINDARAJU et al. 1992), Pinus species using RFLP of 5S RNA spacer region (MORAN et al. 1992), Pseudotsuga species using RFLP of cpDNA, mtDNA and other (STRAUSS et al. 1990), a miocene Taxodium species using rbcL sequence (SOLTIS et al. 1992), seed plants including conifers using rbcL sequence (CHASE et al. 1993, BOUSQUET et al. 1992). However, most studies have used only few genus or rbcL sequence data. Actually, many rbcL sequence data gave us
abundant information of the phylogenetic studies among genera. However, to reveal the species phylogeny, we have to investigate many other genes.

In this study, we focussed on conifer evolution, thus, we tried to construct a physical map of cpDNA of Cryptomeria japonica and then to construct a phylogenetic tree of conifer using site mutation data between several kinds of cpDNA specific genes in conifers.

Materials and Methods

1) Plant materials

Open-pollinated seed of C. japonica D. Don, collected from an individual clone were germinated and grown for 3-6 months in the greenhouse. The needle tissue of these seedlings was used for cpDNA extraction for a physical mapping. Needle tissue were collected from 29 conifer species for phylogenetic relationship study (Table 1).

2) Chloroplast and total DNAs isolation and DNA manipulation

Chloroplast DNA was isolated from needles of C. japonica following by the procedure of OGHARA and TSUNEWAKI (1982). The cpDNA was digested with four restriction endonucleases, namely PstI, SalI, SacI and XhoI. The digested cpDNA was fractionated by 0.7% agarose-gel electrophoresis in TAE buffer (40mM Tris–HCl, 20mM sodium acetate and 2mM EDTA, pH8.0) for estimating of fragment size and/or Southern hybridization. Southern hybridization of sugi cpDNA was carried out with homologous (sugi clones) and heterologous (wheat and tobacco clones) probes to assign the order of restriction fragments of cpDNA. DNA fragments corresponding to the coding region of the genes were prepared after a computer search of restriction site (GENETYX program, SDC Software Development Co. Ltd.). Non-radioactive labeling method and immunological detection of the hybridized fragment was carried out following the protocol of the manufacturer (Boehringer Mannheim Co. Ltd.).

Primer design of specific genes (psbA, rbcL, frxC, psbD and so on) on cpDNA was conducted by DNASIS program, Takara Co. Ltd.. Total DNAs for phylogenetic study were extracted from their needles using a slightly modified CTAB method (MURRAY and THOMPOSON 1980).

3) Polymerase Chain Reaction (PCR) amplification of specific genes on cpDNA and detection of site mutation

PCR amplification of cpDNA genes following procedures: reaction mixture (100μl) contained 10 mM Tris–HCl, pH8.3, 50mM KCl, 1.5mM MgCl2, 0.1% Triton X–100, 0.01% gelation, 0.1mM each dATP, dCTP, dGTP, dTTP, 100 pmol of each primer, 0.5μg
Table 1. Species of conifer examined for \textit{psbA}, \textit{rbcL}, \textit{frxC} and \textit{psbD} restriction site mutation.

<table>
<thead>
<tr>
<th>Species</th>
<th>Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{Cryptomeria japonica} D. Don</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Cryptomeriafortunei} Hoolbrenk</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Taxodium mucronatum} Tenore</td>
<td>BGUT</td>
</tr>
<tr>
<td>\textit{Taxodium ascendens} Bronn</td>
<td>FBRI</td>
</tr>
<tr>
<td>\textit{Taxodium distichum} Rich</td>
<td>FBRI</td>
</tr>
<tr>
<td>\textit{Glyptostrobus pensilis} K. Koch</td>
<td>BGUT</td>
</tr>
<tr>
<td>\textit{Sequoia sempervirens} Endl.</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Metasequoia glyptostroboides} Hu et Cheng</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Cunninghamia lanceolata} Hook.</td>
<td>UFUT</td>
</tr>
<tr>
<td>\textit{Cunninghamia konishii} Hayata</td>
<td>UFCUT</td>
</tr>
<tr>
<td>\textit{Taiwania cryptomerioides} Hayata</td>
<td>UFCUT</td>
</tr>
<tr>
<td>\textit{Sciadopitys verticillata} Seib. et Zucc.</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Chamaecyparis obtusa} (Sieb. et Zucc.) Endl.</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Chamaecyparis pisifera} (Sieb. et Zucc.) Endl.</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Thuja standishii} (Gord.) Carriere</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Thujaos dolabrata} Sieb. et Zucc.</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Taxus cuspidata} Sieb. et Zucc.</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Pinus densiflora} Sieb. et Zucc.</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Pinus thunbergii} Parlatoare</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Pinus koraiensis} Sieb. et Zucc.</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Pinus parviflora} var. \textit{pentaphylla} (Mayr) Henry</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Picea jezoensis} (Sieb. et Zucc.) Carriere</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Picea abies} Karst.</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Abies veitchii} Lindley</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Abies homolepis} Sieb. et Zucc.</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Abies mariesii} Mast.</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Tsuga sieboldii} Carriere</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Pseudesuga japonica} (Shirasawa) Beissner</td>
<td>FFPRI</td>
</tr>
<tr>
<td>\textit{Larix Kaenpferi} (Lamb.) Carriere</td>
<td>FFPRI</td>
</tr>
</tbody>
</table>

FFPRI : Conifer arboretum, Forestry and Forest Products Research Institute, BGUT : Botanical Garden of University of Tokyo, FBRI : Forest Tree Breeding Institute of Forest Agency of Japan, UFCUT : University Forest in Chiba of University of Tokyo, UFUT : University Forest of University of Tsukuba.

DNA, and 2.5 units \textit{Taq} polymerase. Samples were covered with mineral oil. Amplification was carried out using Thermal Temp Controller (Astec Co. Ltd.) for 5 min at 94°C, followed by 35 cycles of 1 min at 94°C, 1 min at 55°C, and 2 min at 72°C, with a final 5 min at 72°C. PCR products were purified by ethanol precipitation and digested by more than ten restriction endonucleases to detect site mutation between specific genes of species. The digested PCR products were put for electrophoresis in 2.0% agarose gel, stained with ethidium bromide, and photographed.
4) Data analysis

Phylogenetic analysis was carried out with PAUP ver. 3.0 (Swofford 1991) using a "branch and bound" algorithm with Wagner parsimony option. The confidence in the phylogenetic tree were examined by 100 replication of bootstrap analysis using the PAVP program.

Results

1) Construction of a physical map and mapping of photosynthesis-related genes

The restriction fragment arrangements were confirmed by homologous hybridiza-

<table>
<thead>
<tr>
<th>Probe (kb) DNA fragment</th>
<th>Restriction enzyme used for digestion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PstI</td>
</tr>
<tr>
<td>S5</td>
<td>P1</td>
</tr>
<tr>
<td></td>
<td>P3</td>
</tr>
<tr>
<td></td>
<td>P13</td>
</tr>
<tr>
<td></td>
<td>P14</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>S7</td>
<td>P1</td>
</tr>
<tr>
<td></td>
<td>PS15</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>S8</td>
<td>P1</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>S9a</td>
<td>P2</td>
</tr>
<tr>
<td></td>
<td>P7</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>S9b</td>
<td>P3</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>S10</td>
<td>P1</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>S11</td>
<td>P3</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>P1</td>
</tr>
<tr>
<td></td>
<td>P7</td>
</tr>
<tr>
<td></td>
<td>P15</td>
</tr>
<tr>
<td></td>
<td>P16</td>
</tr>
<tr>
<td>S12</td>
<td>P10</td>
</tr>
<tr>
<td></td>
<td>P4</td>
</tr>
<tr>
<td></td>
<td>P5</td>
</tr>
<tr>
<td></td>
<td>P9b</td>
</tr>
<tr>
<td>pTB10</td>
<td>P2</td>
</tr>
</tbody>
</table>

NT*: not tested
Fig. 1. Physical map of *C. japonica* in a linear form showing restriction sites of the four endonucleases *PstI*, *SalI*, *SacI* and *XhoI*, and the location of 20 photosynthesis-related genes. The circular DNA has been opened at *SalI* site between S1 (20.4kb) and S9a (8.2kb) (Tsumura et al. 1993).

Table with restriction sites and gene locations.

2) Comparison of cpDNA genome structure between *C. japonica* and Pinaceae species

Gene-mapping data confirm that, like other conifers, the chloroplast genome of *C. japonica* lacks one of the large inverted reperats (Strauss et al. 1988, Lidholm and...
Gustafsson 1991) and is highly rearranged in comparison with other conifers.

3) Phylogenetic relationship among conifers

PCR amplification was checked for 17 genes on cpDNA of 29 conifer species. PCR products for all species were obtained only for six genes. PCR products of two of six genes were not a single band, thus, we detected site mutations to digest PCR products for four genes (psbA, frxC, rbcL, and psbD) (Table 3) by several kinds of restriction endonucleases. Consequently, we found 118 site mutations for four genes in total (Table 4). Taxus was put into the out-group, then, phylogenetic tree was constructed from these data (Fig. 2, Tsumura et al. in preparation). Based on the consensus tree, Cupressaceae and Taxodiaceae except Sciadopitys from a monoplyletic group. Pinaceae also from a monoplyletic group, and genera of this family are separated into two major monophyletic groups, which are Pinus and other genera (Picea, Abies, Tsuga, Pseudotsuga and Larix).

Table 3. Primer sequences for amplification of psbA, frxC, rbcL, and psbD genes.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer</th>
<th>Size</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>psbA</td>
<td>5'-TAGTTCGTGCATAACTTC-3' 5'-CTAGCCTGAAAAACCTT-3'</td>
<td>939bp</td>
<td><em>Pinus contorta</em> (Lindholm and Gustafsson 1991)</td>
</tr>
<tr>
<td></td>
<td>5'-ATAGACATTACGGGAAAGG-3'</td>
<td>779bp</td>
<td><em>Pinus contorta</em> (Lindholm and Gustafsson 1991)</td>
</tr>
<tr>
<td></td>
<td>5'-TGAATAATTCCGATCTGCA-3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5'-TGTCAACAAAAACAGAGAC-3'</td>
<td>1,387bp</td>
<td><em>Pseudotsuga menziesii</em> (Hedrick et al. 1991)</td>
</tr>
<tr>
<td></td>
<td>5'-TTCCATTCACAGCAGC-3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>psbD</td>
<td>5'-TATGACTATACGGCCTTGGTA-3'</td>
<td>1,042bp</td>
<td><em>Nicotiana tabacum</em> (Shinokaki et al. 1986)</td>
</tr>
<tr>
<td></td>
<td>5'-TAGAAACCCTTCAGGGAATA-3'</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Restriction site change of four genes, psbA, frxC, rbcL, and psbD.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Site</th>
<th>Site change</th>
</tr>
</thead>
<tbody>
<tr>
<td>psbA</td>
<td>939bp</td>
<td>23 MspI(2), RsaI(1), Sau3AI(6), TagI(2), MboI(2), SphI(1), HaeII(3), HinII(5), HindIII(1)</td>
</tr>
<tr>
<td>frxC</td>
<td>779bp</td>
<td>21 Sau3AI(1), TagI(5), XbaI(3), SphI(1), SauI(1), RsaI(2), MspI(1), HinII(4), EcoRV(1), HaeII(2)</td>
</tr>
<tr>
<td>rbcL</td>
<td>1,387bp</td>
<td>40 HaeII(7), BglII(3), BamHI(1), MspI(5), PstI(2), RsaI(2), Sau3AI(3), SphI(3), XbaI(1), HhaI(4), TaqI(7)</td>
</tr>
<tr>
<td>psbD</td>
<td>1,042bp</td>
<td>34 HaeII(5), HinII(4), MspI(4), Sau3AI(2), SphI(3), AluI(3), RsaI(2), BsmI(2), HhaI(4), TaqI(5)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>4,147bp</td>
</tr>
</tbody>
</table>


Fig. 2. Strict consensus tree of 29 conifer species based on site change data of psbA, psbD, rbcL and ftrC. Number of each branch indicate the probability of reconstruction of each branch in bootstrap analysis.
Discussion

1) Comparison of cpDNA genome structure of *C. japonica* and species in Pinaceae

Comparing to the physical maps of cpDNA between *C. japonica* (Tsumura et al. 1993) and species in Pinaceae, their gene order is largely differed each other. Structure alternations of the chloroplast genomes of conifers can be traced from the standard type of vascular plant cpDNA. In Pinaceae, the cpDNA genome type of *P. radiata* can be derived from the standard type by six mutations: two deletions and four inversions (Strauss et al. 1988). The cpDNAs of *P. monticola* (White 1990), *P. contorta* (Lidholm and Gustafsson 1991) and *P. thunbergii* (Tsudzuki et al. 1992) have a similar structure to that of radiata pine while that of *Pseudotsuga menziesii* harbors one additional inversion to those of radiata pine (Strauss et al. 1988). Six mutational events, one deletion and five inversions, from the standard cpDNA type are required to produce that of *C. japonica*. But, when the gene order of *C. japonica* cpDNA is compared with that of *Pinus radiata*, the genome structures suggest independent changes so that no simple evolutionary path can be determined.

2) Phylogenetic relationship among genera of Taxodiaceae, Cupressaceae and Pinaceae

Taxodiaceae and Cupressaceae consider to be a monophyletic group but *Sciadopitys* is not included in this group with our results. Hart (1987) also suggested that Taxodiaceae and Cupressaceae from a single monophyletic family from the morphological and chemical traits. Results of immunological similarity of seed proteins (Price and Lowenstein 1989) indicate that a major discontinuity do not exist between Taxodiaceae and Cupressaceae. However, the position of each genus in Taxodiaceae and Cupressaceae are not fully resolved by these data. Taxodiaceae separated into five groups, namely Cryptomeria group with Taxodium and Glyptostrobus, Sequoia group with Metasequoia, Cunninghamia group, and Taiwania.

*Sciadopitys* has been often classified as a morphologically isolated member of the Taxodiaceae or as the separate family Sciadopityaceae Hayata (Hart 1987, Schlarbaum and Tsuchiya 1985). *Sciadopitys* is closely related to Cupressaceae and Taxodiaceae in many morphological traits (Hart and Price 1990). Its chromosome number is n=10, which is apparently branched from that of ancestor of Cupressaceae and Taxodiaceae (n=11) by aneuploid reduction (Schlarbaum and Tsuchiya 1985). Our phylogenetic tree also supported these results.

Pinaceae is also a monophyletic group and separated into two major sub-groups which are *Pinus* and the other genera. Genus *Pinus* is also separated into two groups clearly, which are diploxyon and haploxyon. Hart (1987) also suggested that the
monophyly of the Pinaceae is well established. However, our data have not supported his results completely in the position of each genus.

Summary

To investigate conifer evolution, we constructed a physical map of chloroplast DNA in Cryptomeria japonica. In comparing the cpDNA genome structure between C. japonica and Pinus species, the gene arrangement is largely differed each other. Therefore, we also constructed a molecular phylogenetic tree of 29 conifer species based on site change in four cpDNA genes. Taxodiaceae and Cupressaceae consider to be a monophyletic group but Sciadopitys is not included in this group in our results. Sciadopitys is thought to be branched from an ancestor of Taxodiaceae and Cupressaceae. Pinaceae is also a monophyletic group and separated into two major subgroups which are Pinus and the other genera.

Acknowledgement

We thank Dr. M. SugiuRA for supplying tobacco chloroplast DNA clones. We also thank Mr. K. Yosimura for designing the primers; Drs. N. Tomaru, T. Kondo and K. Tange for collecting materials; Ms. M. Koshiba for help in extracting DNA and PCR amplification. Sincere appreciation is expressed to Dr. T. Kawahara for helpful discussion. This work was supported in part a Grant-in-Aid (No. 02454067) from the Ministry of Education, Science and Culture of Japan and a Grant-in-Aid from Science and Technology Agency of Japan (Encouragement of Basic Research).

References


針葉樹の分子系統進化

津村義彦，大庭喜八郎

〒305 埼玉県朝霞郡前川町松の里1
森林総合研究所

〒305 埼玉県つくば市天王台1-1-1
筑波大学農学系

針葉樹の系統進化を調べるため、スギの葉緑体DNAの物理地図の作成を行った。既報のマツ科のものと比較すると、遺伝子の並びが大きく異なっていた。そこで、さらに針葉樹の系統進化を種レベルで明らかにするため、針葉樹29種について、PCRで増幅した4つの葉緑体DNA上の遺伝子を複数の制限酵素で処理した後、塩基置換を調べた。その結果スギ科とヒノキ科は単系統であると考えられた。しかし分類上問題としてきたコウヤマキはこれらのグループには属しなかった。コウヤマキはスギ科、ヒノキ科を含んだグループの起源種から分岐した種であると考えられた。マツ科もまた単系統であり、マツ属とその他の属に分かれた。

＜質疑応答＞

池谷（果樹試）：研究された方法（特定遺伝子をプローブにしたRFLP法）のデータからは塩基置換頻度が計算できると思うかどうか。
津村：まだやっていません。
藤本（岐阜大）：他殖性であるにもかかわらず、種内の地理的分化が少ないのはなぜか。分布している環境の差が小さいのではないか。他殖性牧草で、世界的に分布の広い種では、種内の地理的分化は大きい。
津村：例えばスギでは、風媒花なのでかなり遠くまで花粉を飛ばせるため、近隣の集団間での分化は少ない。牧草の場合は花粉をそれほど飛ばせないのでないか。
武田（東京大）：針葉樹では遺伝的変異性が高いそうですが、そのために地域分化の程度の推定や、分子系統進化での系統関係の推定が不安定になることはないでしょうか。
津村：葉緑体遺伝子のconserveされた領域を使っているので、問題はないと考えている。
座長（近藤）：このような手法は他に使われているか。
津村：ある草本植物で、rpoCの4.2kbを増やしてその間でのRFLPを使って属間での変異を見ている。
RFLP PATTERNS DETECTED BY THE TRANSPOSONS, MAG AND K1.4, IN SILKWORM RACES

Toshiki TAMURA, Toshio KANDA, Kenji YUKUIRO, Yuji YASUKOCHI, Norihide HINOMOTO, Kunimitsu SHIMIZU, Eiichi KOSAGAWA and Teruo OKAJIMA

Institute of Sericultural and Entomological Science.
Owashi 1-2, Tsukuba-shi, Ibaraki-ken, 305

Introduction

The domesticated silkworm, *Bombyx mori*, is believed to have originated in ancient China and distributed throughout the countries of the world. Since sericulture was a very important industry in Japan, many races present in different countries were collected from almost the whole world from middle of Meiji to the beginning of Showa eras (1910–1950). The breeding of silkworm races for commercial purposes in these eras was performed by utilizing these races, and economically important characters were very quickly improved. Stocks of more than 1000 races, including geographical, improved, and mutant races, are now maintained in Japanese Institutes and Universities, for example, NISES, Kyushu University, and so on (IYAMA et al., 1988; SORITA, 1991). In addition, many races are maintained in China, India, Korea and European countries. Developing new methods to determine the genetic relationships among these races is of great interest from the viewpoint of silkworm origin and evolution, as well as for application to practical breeding.

In the silkworm, many genes, for example, fibroin, sericins, chorion, etc., are already cloned and their physical structures are determined (SUZUKI et al., 1990; GOLDSMITH and KAFATOS, 1984). Furthermore, clones showing polymorphic patterns in Southern blotting are screened from cDNA or genomic DNA libraries (GOLDSMITH and SHI, 1991; YUKUIRO et al., 1992), and the method of random amplified polymorphic DNA is also applicable to silkworms (SHIMADA et al., 1991). Thus, there are many candidates to study differences on the level of DNA polymorphism in silkworm stocks and races. We chose the transposons mag and k1.4 as candidates for suitable marker genes to detect DNA polymorphism because the copy number of these genes is around 10 (MICHAILLE et al., 1990; UEDA et al., 1986), which is in an appropriate range for characterizing stock differences. In addition, DNA polymorphism could appear even
among genetically similar races because these transposons are thought to give more polymorphic bands than conventional genes.

In this communication we examined RFLP patterns of mag and k1.4 in typical geographical, improved and commercial parent races. The results showed that the patterns obtained by these transposon DNAs are very useful to characterize genetic stocks and races.

**Materials and Methods**

1) **Silkworm races used.**

Eleven geographical and 3 improved stocks, 8 commercial races and 2 mutant races were used to compare RFLP patterns detected by the two transposons (Table 1).

<table>
<thead>
<tr>
<th>Table 1. Races used for the experiment.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geographical Tropical Race</td>
</tr>
<tr>
<td>Geographical Japanese Race</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Geographical Chinese Race</td>
</tr>
<tr>
<td>Geographical European Race</td>
</tr>
<tr>
<td>Geographical Korean Race</td>
</tr>
<tr>
<td>Improved Japanese Race</td>
</tr>
<tr>
<td>Improved Chinese Race</td>
</tr>
<tr>
<td>Commercial Japanese Parent Race</td>
</tr>
<tr>
<td>Commercial Chinese Parent Races</td>
</tr>
<tr>
<td>Mutant Strain</td>
</tr>
</tbody>
</table>

The 11 races were from the Laboratory of Genetic Resources, Department of Insect Genetics and Breeding, National Institute of Sericultural and Entomological Science (NISES). Two, pnd-p^a and DT, were from the Laboratory of Gene Engineering of the same department. Two parent races, J01 and C01, were from the Laboratory of Silkworm Egg Production, Department of Sericulture, NISES. The eight races, Asa, Hi, Fu, Yoh, Toh, Kai, Kouhaku and Hei, were obtained courtesy of the Institute of Silkworm Genetics and Breeding. The standard races, which were chosen from a report of the Japanese Sericultural Science Society (1986), were mainly used in this experiment.

2) **Isolation of Insect DNA**
Total DNA was isolated from silk glands. Around 50 larvae of each race were reared on artificial diet or fresh mulberry leaves. When the larvae attained 2 to 4 days of development after the final ecdysis, the posterior portion of the silk glands was collected and washed twice in 1 x SSC. Then, the silk glands were frozen by liquid nitrogen and stored at −80°C. The glands were thawed in DNA extraction buffer (50mM Tris pH 8.0/100mM NaCl/20mM EDTA) containing pronase K (150 μg/ml) and homogenized with a dounce homogenizer. Ten percent of SDS solution was added to a final concentration of 1% SDS and the homogenate was incubated at 50°C for 2 h. The solution was then extracted three times with phenol, and three times with ether, and then subjected to EtBr-CsCl₂ density gradient ultra centrifugation. After removing EtBr with isopropanol saturated with CsCl₂-saturated DW, DNA recovered from the centrifugation was dialyzed against TE.

3) Preparation of probes

A plasmid containing the retrotransposon mag sequence and an M13 phage with the transposable element k1.4 were kindly obtained from Dr. A. Garel and Dr. H. Ueda, respectively. To prepare a fragment with a suitable region of mag, two primers were synthesized and the most conserved region was amplified by PCR using the plasmid or silk worm genomic DNA (Fig. 1a). M13 phage DNA was purified by the alkali method and EtBr-CsCl₂ gradient centrifugation. The k1.4 insert of the purified phage DNA was cut out by digestion with restriction enzymes and used as a probe template. Both DNA templates of mag and k1.4 were labeled with random primer using the DIG system (BM) or [α-³²P] dCTP (Amersham).

4) Blotting and hybridization.

Genomic DNA was digested with restriction enzymes (NEB) and 1 μg of digested DNA (3 μg for DIG detection system) was run on an 0.8% agarose gel for around 3 h at 80 V in TPE buffer. The DNA was transferred to a nylon membrane (Hybond N, Amersham) by vacuum blotting (VacuGene™, Pharmacia LKB), and fixed with UV light using a Funa-UV-Linker (Funakoshi). Hybridization was performed using 50% formamide and Denhart solution or the DIG kit solution. Washing was carried out twice at room temperature with 2 x SSC/0.5% SDS for 15 min and twice at 65°C with 0.1 x SSC/0.5% SDS. The washed membrane was exposed to X-ray film (X'Omat, Kodak) for 1–3 days.

Results

1) RFLP patterns detected by the retrotransposon mag.
Fig. 1. Physical map of retrotransposon mag (a) and RFLP pattern (b). Pvu, Kpn and P+K indicate that DNAs were digested with Pvu II, Kpn I and both enzymes, respectively. Lane 1, C108; lane 2, Daizo; lane 3, Cambodge; lane 4, F1 between C108 and Daizo; lane 5, F1 between C108 and Cambodge. Bar represents region used as the probe. Labelling and hybridization were performed by DIG kit (BM).
Mag, which is shown in Fig. 1a, is a retrotransposon originally found in the upstream region of a sericin gene. It consists of approximately 4.5 kb DNA containing two direct terminal repeats of 77 bp and two open reading frames (Michaille et al., 1990). The recognition sites for one of two restriction enzymes, Kpn I and Pvu II, are present near each end. We first examined whether an RFLP pattern suitable for the estimation of genetic relationships could be obtained using mag as a probe. The results obtained with three representative races and their F₁ hybrids are shown in Fig. 1b. As we expected, the complexity of banding patterns was markedly different by the digestion with a single enzyme or double digestion with the two enzymes which possess a recognition site near the 5' or 3' ends of the transposon sequence. When digested with just one of the enzymes, one site appearing in a band must come from outside the mag sequence, while one site must be in the transposon. Thus, a polymorphic band observed under these conditions is thought to be due to a positional difference of the inserting site or an evolutionary change in a recognition site. On the other hand, polymorphic bands caused by structural changes in mag itself or sequential changes in the enzyme recognition sites would appear after double digestion using the two enzymes. Therefore, we predicted that a highly complex pattern would be observed when using just one enzyme, and a much simpler pattern would be obtained after treatment with the two enzymes, Kpn I and Pvu II. The results shown in Fig 1b showed that our prediction was correct, and that we can adjust the complexity of the RFLP pattern of mag by choosing suitable restriction enzymes. In addition, from the comparison of RFLP patterns of three races and their F₁s we confirmed that the polymorphic bands detected by mag were stably transferred to the next generation.

To know how to apply this method efficiently to other races, the RFLP patterns were studied in a greater number of races. We used Kpn I and Pvu II double digested DNAs for this purpose because we were afraid that the extremely high complexity appearing in a single enzyme digest would make it impossible to analyze the RFLP patterns, especially when comparing many geographical races among different filter membranes. Representative RFLP patterns obtained in this condition are shown in Fig. 2. A total of fifteen bands was detected, and all of them could be assigned as present or absent without ambiguity. Eleven of these bands were polymorphic among the races, and four were ubiquitous. The presence or absence of these bands are scored in Table 2. The divergence between geographical races was rather high, whereas only small numbers of polymorphic bands were detected within the Japanese or Chinese commercial parent races. This means that the RFLP pattern detected by the double digestion of Kpn I and Pvu II was suitable to discriminate races with rather large genetic distance, such as among the geographical races, but not for genetically close races, like within the commercial races.
Fig. 2. RFLP pattern detected by mag. The central region of mag (bar in Fig 1a) were labelled by random primer extension using \((a^{-32}P)\) dCTP. Genomic DNAs of each stocks and races were digested with Pvu II and Kpn I. Lane 1, Mysore; lane 2, Cambodge; lane 3, Akajuku; lane 4, Nihonnishiki; lane 5, Koishimaru; lane 6, J115; lane 7, J124; lane 8, J01; lane 9, Daizo; lane 10, Kousetsu; lane 11, Kansen; lane 12, C108; lane 13, C01 and lane 14, E16.
Table 2. Analysis of RFLP pattern detected by mag.

<table>
<thead>
<tr>
<th>races</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
<th>i</th>
<th>j</th>
<th>k</th>
<th>l</th>
<th>m</th>
<th>n</th>
<th>o</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mysore</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Cambodge</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Akajuku</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Nihonnishiki</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Koishimaru</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>J115</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>J124</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>J101</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Daizo</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Kousetu</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Kansan</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>C108</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>C01</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>E16</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Kansanmin</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>pnd, pS</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>C108xCambodge</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>C108xDaizo</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Asa</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Hi</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Fu</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Yoh</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Toh</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Kai</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Hei</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

To obtain more polymorphic bands within commercial races, single enzyme digestion was performed with EcoRI, and the patterns obtained are shown in Fig. 3. This digest apparently gave greater numbers of polymorphic bands; only two of 17 bands were ubiquitous, and others showed racial differences. Even within the Japanese or Chinese parent races, four polymorphic bands could be determined in each cases. This suggested that a single enzyme digestion with EcoRI was superior for discriminating among races with similar genetic background.

2) RFLP pattern obtained with the transposable element k1.4.

In order to ask if similar results obtained with mag could also be found using the transposable element k1.4, we performed Southern blot analysis using the k1.4 sequence as a probe against DNAs of the same three races and their F1's. Unlike mag, k1.4, whose size is around 1.4kb, is a transposable element with a 12 bp inverted repeat at both ends (UEDA et al., 1986). Hae III sites are present at both ends (Fig. 4a), and
Fig. 3. RFLP pattern detected by mag. The DNAs were digested with EcoRI. Lane 1, C108; lane 2, Daizo; lane 3, Cambodge; lane 4, F1 between C108 and Cambodge; lane 5, F1 between C108 and Daizo; lane 6, Asa; lane 7, Hi; lane 8, Fu; lane 9, Yoh; lane 10, Toh; lane 11, Kai; lane 12, Kouhaku; lane 13, Hei; lane 14, pnd -p4 and lane 15, DT.

could be used as suitable sites to produce an RFLP pattern. We first compared the RFLP patterns obtained by digestion with Hae III or with one of two other enzymes, Fsp1 and BamH1, both of which have no recognition sites within the k1.4 transposon sequence. As shown in Fig. 4, digestion by Fsp1 and BamH1 gave larger numbers of bands compared to that of Hae III. A comparison of patterns produced by digestion with Fsp1 or BamH1 indicated that it was possible to determine conclusively the presence or absence of relatively few bands because of their complexity of patterns. On the other hand, the presence or absence of bands larger than 1.4kb that were produced by Hae III could be clearly discriminated in different races. Therefore, we used RFLPs of Hae III-digested DNA for further analysis as the representative patterns obtained with this transposon. The patterns of 17 races were scored; it was possible to use 8 bands for analysis, of which six were polymorphic among the races.
Fig. 4. Physical map of transposable element k1.4 (a) and RFLP pattern. Five different genomic DNAs were digested with Hae III (Hae), Fsp I (Fsp) and BamHI (Bam). Lane 1, C108; lane 2, Daizo; lane 3, Cambodge; lane 4, F1 between C108 and Daizo and lane 5, F1 between C108 and Cambodge.
3) Estimation of genetic similarity among races

Since the presence or absence of a total 23 bands were determined among the 17 races, we calculated their genetic similarity by the following formula: genetic similarity (GS) = numbers of shared bands between two races/numbers of total bands observed in the two races. Although the total numbers of bands were somewhat small to apply to this method, we thought it would be useful to calculate this value because no such trials have yet been reported in the silkworm. The values obtained are shown in Table 3, and appear to be in very good agreement with known historical relationships of the races. For example, the value between two tropical races, Cambodge and Mysore, was 0.96, and those among Japanese geographical races, Akajuku, Nihonnishiki and Koishimaru, were more than 0.92. On the other hand, values between Japanese and Chinese races were low compared to those within Japanese races or Chinese races, although there were some exceptional cases. Based on these genetic similarities among the races, a genetic association was calculated by the UPGMA method (Nii, 1987). As shown in Fig. 5, the genetic association calculated from GS agreed very well with the genetic relationships among the races.

**Table 3. Genetic similarity obtained from RFLP pattern of 17 different races.**

<table>
<thead>
<tr>
<th>Races</th>
<th>Mys</th>
<th>Cam</th>
<th>Aka</th>
<th>Nih</th>
<th>Koi</th>
<th>J115</th>
<th>J124</th>
<th>J01</th>
<th>Dai</th>
<th>Koe</th>
<th>C108</th>
<th>C01</th>
<th>El6</th>
<th>Ksa</th>
<th>Asa</th>
<th>Toh</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mysore (Mys)</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cambodge (Cam)</td>
<td>0.92</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akajuku (Aka)</td>
<td>0.92</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nihonnishiki (Nih)</td>
<td>0.92</td>
<td>0.96</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koishimaru (Koi)</td>
<td>0.92</td>
<td>0.96</td>
<td>0.92</td>
<td>0.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J115</td>
<td>0.87</td>
<td>0.91</td>
<td>0.87</td>
<td>0.87</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J124</td>
<td>0.78</td>
<td>0.82</td>
<td>0.78</td>
<td>0.78</td>
<td>0.87</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J01</td>
<td>0.83</td>
<td>0.87</td>
<td>0.92</td>
<td>0.92</td>
<td>0.87</td>
<td>0.87</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taiyo (Dai)</td>
<td>0.75</td>
<td>0.78</td>
<td>0.75</td>
<td>0.75</td>
<td>0.83</td>
<td>0.78</td>
<td>0.78</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kousetzu (Koe)</td>
<td>0.77</td>
<td>0.80</td>
<td>0.84</td>
<td>0.84</td>
<td>0.80</td>
<td>0.72</td>
<td>0.77</td>
<td>0.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kassen (Koe)</td>
<td>0.75</td>
<td>0.78</td>
<td>0.75</td>
<td>0.75</td>
<td>0.70</td>
<td>0.70</td>
<td>0.67</td>
<td>0.75</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C108</td>
<td>0.78</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.87</td>
<td>0.70</td>
<td>0.64</td>
<td>0.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C01</td>
<td>0.85</td>
<td>0.80</td>
<td>0.77</td>
<td>0.69</td>
<td>0.77</td>
<td>0.80</td>
<td>0.80</td>
<td>0.92</td>
<td>0.69</td>
<td>0.71</td>
<td>0.62</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>El6</td>
<td>0.77</td>
<td>0.80</td>
<td>0.77</td>
<td>0.77</td>
<td>0.80</td>
<td>0.80</td>
<td>0.92</td>
<td>0.77</td>
<td>0.71</td>
<td>0.68</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kansai (Ksa)</td>
<td>0.67</td>
<td>0.61</td>
<td>0.75</td>
<td>0.75</td>
<td>0.67</td>
<td>0.61</td>
<td>0.61</td>
<td>0.67</td>
<td>0.58</td>
<td>0.69</td>
<td>0.75</td>
<td>0.61</td>
<td>0.67</td>
<td>0.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asa</td>
<td>0.87</td>
<td>0.91</td>
<td>0.87</td>
<td>0.87</td>
<td>0.96</td>
<td>0.91</td>
<td>0.91</td>
<td>0.78</td>
<td>0.78</td>
<td>0.80</td>
<td>0.70</td>
<td>0.73</td>
<td>0.72</td>
<td>0.72</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>Toh</td>
<td>0.88</td>
<td>0.83</td>
<td>0.80</td>
<td>0.80</td>
<td>0.80</td>
<td>0.75</td>
<td>0.83</td>
<td>0.88</td>
<td>0.72</td>
<td>0.74</td>
<td>0.72</td>
<td>0.83</td>
<td>0.89</td>
<td>0.81</td>
<td>0.72</td>
<td>0.83</td>
</tr>
</tbody>
</table>

**Discussion**

We confirmed that RFLP patterns suitable for studying genetic relationships
among the domesticated silkworm races can be obtained by using the transposons, mag and k1.4. From the RFLP patterns, at least 25 bands were assigned to be used for further analysis. The comparison of banding patterns in different races showed that eighteen of them were polymorphic, and genetic similarity among the races were scored from the banding patterns. A genetic association among the races was also calculated from the values of genetic similarity. The values and association showed
that these results obtained from the RFLP patterns agree very well with the historical localization of geographical races. That is, for geographical Japanese, Chinese and European races, each of them fell into the same group.

No other similar experiments of measuring genetic relationships among silkworm races using DNA polymorphism have been published to date. Most works to determine genetic relationships have been carried out only by the use of morphological difference or protein polymorphism, including isozymes. However, the total amount of information obtained with isozymes, like phosphatase, esterase, etc., or blood proteins, is rather limited. For example, one of the most reliable sets of data, which reports on blood protein patterns using PAGE by GAMO and OHTUKA (1980), demonstrates that only 11 protein bands can be used for this analysis. Since the numbers of polymorphic bands can easily be increased in our method by changing the enzyme for DNA digestion or using other DNA sequences as probes, this RFLP method is apparently superior to the earlier methods. However, the data reported in this communication was only obtained with limited numbers of races and clones. So, it is necessary to examine a greater number of other races and clones for further developing this method, as well as extending it to the application of PCR.

From the analysis of isozymes and blood proteins (YOSHITAKE, 1968; GAMO and OHTUKA, 1980), the domesticated silkworm is hypothesized to have originated from Chinese univoltine races. One reason for this idea is the fact that the highest genetic diversity is observed in these Chinese races. A similar phenomenon was also observed in our RFLP analysis. The genetic similarity within the geographical Chinese races was lower than within other geographical races, indicating that our results using transposable elements are consistent with this hypothesis. We expect that experiments using increased numbers of genes and races will give a more clear conclusion about the origin and distribution of the geographical races.

Our data also indicated that the RFLP patterns obtained with transposons like mag and k1.4, might be useful to discriminate and characterize commercial parent races. In the domesticated silkworm, morphological differences are very small, especially within Japanese or Chinese parent races, in which differences are mainly due to economical characters, like cocoon shell weight, viability, and cocoon size. Because of their polygenic natures, these characters are rather unstable, so it is almost impossible to identify races once they are lost or confused. The method reported here may provide a way to identify such races by names by analyzing RFLP patterns. In addition, the genetic similarity of two commercial parent races can easily be calculated from the RFLP patterns. Knowing the similarity will contribute to finding out the most suitable combination between Chinese and Japanese parent races for making hybrids. The similarity value is also thought to be useful to predict the best cross as
a starting strain for breeding a new commercial parent race. We hope that our RFLP method and other procedures for detecting DNA polymorphism in the domesticated silkworm will be introduced into practical breeding and lead to the development of a new breeding method for the silkworm.

Summary

We first examined whether an appropriate RFLP pattern for studying the genetic relationships among silkworm races can be obtained by using DNA sequences of the transposons, mag and k1.4, as probes. Comparison of the RFLP patterns among the representative geographical races showed that highly polymorphic bands appeared in these sequences when digested by restriction enzymes which recognized both ends of the transposons. A total of 25 bands were observed in the patterns, of which 18 were polymorphic. Genetic similarity and association among the races used were calculated from the patterns. The values obtained agreed very well with the relationships from historical descriptions.

Acknowledgments

We thank for Dr. Marian R. GOLDSMITH for her useful suggestions and comments on the manuscript.

References


蚕の品種間に見られるトランスポゾンのRFLP

田村俊樹、神田俊男、行弘研二、安河内祐二
日本典秀、清水久仁光、小瀬川英一、岡島輝男
農林水産省畜系・昆虫農業技術研究所
〒305 茨城県つくば市大和1-2

蚕は中国を発祥の地として、世界中に伝播した昆虫であると考えられている。日本では、明治から昭和の始めにかけて、世界各国の在来品種が収集され、今日では1000以上もの系統や品種が保存されている。これらの品種の遺伝的な類縁関係を調べる手法を開発することは、品種の改良などの実用的な意味だけではなく、蚕の起源や進化を研究する上でも重要である。本研究では既にクローニングされ、構造の明らかにされている2種のトランスポゾンmagとk1.4のDNAを利用して、蚕の品種間でどの程度のRFLPが検出できるかを調べた。

magはレトロタイプで両末端に77bpのダイレクトリピートと2つのオープンリーディングフレームを持つ、約4.5kbのトランスポゾンである。この遺伝子の塩基配列の一部を用いて、どの程度の多型が検出できるかを調べたところ、一度に10数種類のバンドが検出され、しかもこれらのバンドの品種間差は大きかった。さらに、観察された多型は次世代に安定的に遺伝することが分かった。得られるバンドの形や多型の頻度は用いる制限酵素の種類によって変わり、遺伝的に大きく異なっている品種間の違いを見る場合と非常に近い関係にある品種間の差を検出する場合とで使い分けすることが可能であった。また、末端に12bpのインパラーテッドリピートを有する1.4kbのトランスポゾンk1.4のクローニングを用いた場合も、magを用いた場合と殆ど同じ結果が得られ、両者とも品種間のRFLPパターンを比較するためには、有効で効率的な遺伝子であることが明らかになった。そこで、日本種、中国種、欧州種、熱帯種などからこのうちの代表的な15品種並びに8種類の実用品種を選び、これらの品種において検出されるRFLPパターンを調べた。その結果、明確に品種間で比較できるバンドとして25種類が同定され、これらの18種が品種間で多型があり、品種の特徴をバンドの有無から数値化することができた。この値に基づいて、品種間の遺伝的な類似性を計測した結果、従来から遺伝的に近いと考えられている品種間では相似性は高く、離れているもののでは低いことから、今回用いたトランスポゾンのRFLPバターンから得られる品種間の遺伝的な類縁関係は、これまでの品種に関する知見と良く一致すると判断された。さらに、品種間で得られた相似性の値を利用して、UPGMA法による各品種のグループ化を試みた結果、各品種は各地域毎に分類され、この点についても従来の知見と矛盾しないことが明らかにされた。また、地域品種毎に遺伝的な相似性を比較した場合、中国種間では低く、遺伝的変異は中国種において最も大きいことが分かった。これに
とは、蚕の起源が中国であるとの説を強く支持するものである。

＜質疑応答＞
黒田（生物研）：カイコの類縁関係については、これまでアイソザイムでの優れた業績があるが、現時点で結果に相違がみられますか。
田村：デンプングラムを書く段階になっていないのでまだ分からないが、この方法の方がアイソザイムより多くの情報が得られています。
黒田：今回の実験に使われたDNAはどこから採られたのか。
田村：構造段階から採っています。
森田（茨城大）：品種の遺伝的なhomogeneityはどの程度か。品種内の変異についてはどうか。
田村：集団で保存されているので、それほど純粋ではない。
座長（鶴生）：カイコは植物で言えば他花授精なので、inbred depressionが起こる。系統は50個体の蛾から生じる個体を混ぜて維持している。
CHLOROPLAST DNA DIVERSITY IN *PYRUS*
AND RELATED GENERA

Hiroyuki *Iketani*
Fruit Tree Research Station
2-1, Fujimoto, Tsukuba, Ibaraki, 305

Introduction

The useful characters of the cultivated crops are the results between human selection and evolution. Recently, molecular biology has made possible to analyze the complicated characters such as morphological ones. In the near future, evolution of these characters will be able to be understood (CHASAN 1993). On the other hand, before the evolutionary analysis of the characters of some organisms, phylogenetic pattern of them should be elucidated. For this purpose, the methods independent from the analyzed characters are desirable. Molecular characters, especially chloroplast DNA (cpDNA) fulfill this condition and they are widely used for phylogenetic analyses in plants (PALMER et al. 1988, CRAWFORD 1990).

*Pyrus* and related genera contain several important temperate fruit trees such as pear, apple, and quince. They belong to the subfamily Maloideae (Rosaceae) and share many characters, still they differ in several important characters such as fruit ones (IKETANI and OHASHI 1991). However, the systematics of this group still remains many problems and the relationships among genera are little known because of several reasons, e. g., lack of "good" synapomorphic characters, incongruence between morphological diversity and high cross-ability among taxa, and so on.

We analyzed the chloroplast DNA diversity of these genera with restriction fragment length polymorphisms (RFLPs) which are widely used in comparison of interspecific or intergeneric taxa of seed plants.

**Materials and Methods**

1) DNA extraction

Four accessions of *Pyrus* and four accessions belonging to related three genera were used as materials (Table 1). About 10 g of the fresh leaves were powdered with
Table 1. Materials investigated.

<table>
<thead>
<tr>
<th>No.</th>
<th>Species</th>
<th>cultivar name</th>
<th>location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Pyrus pyrifolia Nakai</td>
<td>‘Nijisseiki’</td>
<td>FTRS</td>
</tr>
<tr>
<td>2.</td>
<td>P. ussuriensis Maxim.</td>
<td>‘Iwate No. 1’</td>
<td>FTRS</td>
</tr>
<tr>
<td>3.</td>
<td>P. betulaefolia Deane.</td>
<td>‘Karamamashii’</td>
<td>FTRS</td>
</tr>
<tr>
<td>4.</td>
<td>P. communis L.</td>
<td>‘La France’</td>
<td>FTRS</td>
</tr>
<tr>
<td>5.</td>
<td>Maltus domestica Borkh.</td>
<td>‘Fuji’</td>
<td>FTRS</td>
</tr>
<tr>
<td>6.</td>
<td>M. tschonoskii Maxim.</td>
<td>—</td>
<td>Miyagi</td>
</tr>
<tr>
<td>7.</td>
<td>Cydonia oblonga Mill.</td>
<td>‘Champion’</td>
<td>FTRS</td>
</tr>
<tr>
<td>8.</td>
<td>Chaenomeles sinensis Koehne</td>
<td>—</td>
<td>NSM</td>
</tr>
</tbody>
</table>

1) FTRS: cultivated at Fruit Tree Research Station, Miyagi: wild tree in Sendai, Miyagi Pref., NSM: cultivated at Tsukuba Botanical Garden, National Science Museum.
2) No cultivar name

motor and pestle under liquid nitrogen. They were suspended in 500 ml of the wash buffer (50 mM Tris-HCl pH8.0, 5 mM MgCl₂, 0.3 M sucrose, 0.2 % 2-Mercaptoethanol, 1 % PVP-400) and filtered with three layers of cheese cloth and three layers of miracloth. The filtrate was centrifuged at maximum 1,000 × g, for 10 minutes, at 4°C. Then, the pellet was suspended with 10 ml of lysis buffer (50 mM Tris-HCl pH 8.0, 20 mM EDTA, 1 % Sarcosine). At last, DNA was purified with CsCl-ultracentrifuge.

2) Southern hybridization and detection of the polymorphisms

About one μg of the purified DNA was digested with the following 16 different restriction endonucleases, respectively: *Apa* I, *Bam* H I, *Dra* I, *Eco* O109 I, *Eco* R I, *Eco* R V, *Hind* III, *Kpn* I, *Pst* I, *Pvu* II, *Sac* I, *Sal* I, *Sma* I, *Sty* I, *Xba* I, and *Xho* I. DNA fragments were separated on 0.8 % agarose gels, then transferred to nylon membranes (Hybond N, Amersham). *Sedum album* cpDNA clones (HAM et al. 1992) were used as the hybridization probes. 22 different clones representing 99 % of the chloroplast genome were divided into six groups and each group of clones was used as a single probe. Clones were mixed as follows: No. 1 = S4+S7, No. 2 = P6+P8+S14+S17+SP3, No. 3 = S5+S13+S16+SP4, No. 4 = P10+SP1+SP2, No. 5 = S9+S10+S12+S15, No. 6 = H7+H8+H9+S8. They were labeled with Digoxigenin-11-dUTP using DNA Labeling Kit (Boeringer). Hybridization was performed at 65°C, overnight. Then, membranes were washed twice with 2×SSC-0.1% SDS, for five minutes, at room temperature and washed twice with 0.2×SSC-0.1% SDS, for 15 minutes, at 65°C. Finally they were detected with chemiluminescence using DIG Luminescent Detection Kit (Boeringer).

3) Data analysis
From the obtained hybridization band patterns, only site mutations were estimated and used for data analysis, because the length mutations are difficult to identify homology and evolutionary polarity as well as statistic nature is unknown. Estimate of the number of nucleotide substitutions per site between each of the cpDNAs were calculated by the NEI and Li (1979)'s method using the data from 14 enzymes (excluding EcoO109 I and Sst I: they recognizes multiple sites). A Wagner parsimony phylogenetic tree was constructed with the MIX program of the Phylogenetic Inference Package (PHYLIP) ver. 3.0 created by J. Felsenstein (Washington University, U. S.A.). Bootstrap reliability was estimated with the BOOT program of the same package.

Results and Discussion

1) Restriction site polymorphisms

The fragments visualized for each enzyme with each probe were corrected for the fact that adjacent probes were hybridized to the same fragment. It was estimated that a total of 282 sites had been assayed. Of the total, 13 were polymorphic due to site mutations (Table 2). Among them, six (mutation 8–13 in Table 2) were autoapomor-

<table>
<thead>
<tr>
<th>Mutation No.</th>
<th>Enzyme</th>
<th>Probe No.</th>
<th>Mutation</th>
<th>Material No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>BamH I</td>
<td>4</td>
<td>4.1 + 2.5 + 1.6</td>
<td>5, 6</td>
</tr>
<tr>
<td>2.</td>
<td>BamH I</td>
<td>6</td>
<td>6.3 + 5.2 + 1.1</td>
<td>1, 2, 3, 4</td>
</tr>
<tr>
<td>3.</td>
<td>EcoO109 I</td>
<td>1</td>
<td>6.4 + 5.6 + 1.2</td>
<td>1, 3</td>
</tr>
<tr>
<td>4.</td>
<td>EcoO109 I</td>
<td>3</td>
<td>3.9 + 2.2 + 1.7</td>
<td>5, 6</td>
</tr>
<tr>
<td>5.</td>
<td>Kpn I</td>
<td>4</td>
<td>16 + 13 + 29</td>
<td>5, 6</td>
</tr>
<tr>
<td>6.</td>
<td>Sal I</td>
<td>4</td>
<td>23 + 17 + 6</td>
<td>6, 9</td>
</tr>
<tr>
<td>7.</td>
<td>Xba I</td>
<td>4</td>
<td>10.2 + 8.6 + 1.6</td>
<td>1, 3</td>
</tr>
<tr>
<td>8.</td>
<td>Apa I</td>
<td>4</td>
<td>23 + 12 + 11</td>
<td>6</td>
</tr>
<tr>
<td>9.</td>
<td>Pvu II</td>
<td>4</td>
<td>10.8 + 8 + 2.8</td>
<td>8</td>
</tr>
<tr>
<td>10.</td>
<td>Sac I</td>
<td>3</td>
<td>4.2 + 2 + 6.2</td>
<td>8</td>
</tr>
<tr>
<td>11.</td>
<td>Sal I</td>
<td>6</td>
<td>23 + 15 + 8</td>
<td>1</td>
</tr>
<tr>
<td>12.</td>
<td>Xba I</td>
<td>6</td>
<td>4.6 + 2.4 + 2.2</td>
<td>9</td>
</tr>
<tr>
<td>13.</td>
<td>Xho I</td>
<td>4</td>
<td>7.5 + 6.6 + 0.9</td>
<td>9</td>
</tr>
</tbody>
</table>

phic mutations, that is, only one material differed from the others. Calculated nucleotide diversities ranged from 0 to 0.007. These estimates are similar to or lower than values obtained from the previous studies with closely related species (RIESEBERG et al. 1988; DALLY and SECOND 1990; DOYLE et al. 1990a; DOEBLEY et al. 1992). Thus,
they are considerable low for intergeneric comparison. Genetic diversities of taxa of Maloideae have been said to be lower than those of other plants because of their relative facility of interspecific or intergeneric cross-ability and graft-compatibility (Kovanda 1965). The result supported this hypothesis and except for a preliminary report by Dickson et al. (1989), it is the first demonstrating evidence on intergeneric comparison, following the isozyme study of Dickson and Weeden (1991) on interspecific comparison on Malus.

2) Phylogenetic relationships

Only one cladogram was obtained from Wagner parsimony analysis (Fig. 1). In

![Inferred cladogram](image)

Fig. 1. Inferred cladogram. Solid bars indicate synapomorphic mutations. Open bars indicate homoplasious mutations. The numbers on the bars correspond mutation numbers in Table 2. Numbers left of the branches indicate bootstrap reliability per 1000.

this cladogram, both Pyrus and Malus were combined into a clade, respectively, although the bootstrap reliability of the former was not high. Traditionally, these two genera have been thought to the most closely related among the genera of Maloideae (Hutchinson 1965). However, the obtained cladogram showed more closely relationships between Pyrus and Cydonia, though the monophyly of these two was not demonstrated. For detailed discussion of the phylogenetic relationship, more statistically
reliable data are needed, that is, a larger number of restriction enzymes are needed for investigation.

Three mutations were appeared in *Pyrus* species. One (No. 11 in Table 2) was found only in ‘Nijisseiki’ (*P. pyrifolia*) and two (No. 3 and 7 in Table 2) were appeared in ‘Nijisseiki’ and ‘Karamamenashi’ (*P. betulaefolia*). Morphologically, *P. pyrifolia* is much more closer to *P. ussuriensis* than *P. betulaefolia*. Two reasons can be thought for this incongruence. One is that the morphological resemblance of *P. pyrifolia* and *P. ussuriensis* is due to parallel evolution or mere symplesiomorphy. The other is that it is due to introgression between *P. pyrifolia* and *P. betulaefolia*. Similar phenomena due to interspecific introgression are reported (Doyle et al. 1990b; Arnold et al. 1991, Whittmore and Schaal 1991). And considering the high cross-ability and frequent appearances of natural hybrids between Asiatic species of *Pyrus*, the latter is more plausible. To clarify this hypothesis, more detailed studies on the distribution of polymorphisms with many other cultivars and strains of Asiatic pears should be needed. And studies on the nuclear genome are also important. Through these works, the evolutionary process of the character diversification of *Pyrus* will be become more clear, and it may be helpful for the analysis of useful characters.

**Summary**

Restriction fragment length polymorphisms of chloroplast DNA in *Pyrus* and related genera were investigated. Observed site mutations were quite a few and estimated nucleotide diversities were low. This supported the hypothesis that the genetic diversities between the genera of Maloideae are low as expected from the high cross-ability between genera.

**Acknowledgements**

The author would like thank to Drs. T. Hayashi and N. Matsuta of Fruit Tree Research Station for their valuable advises, to Mr. T. Manabe of Meiji University for technical assistance, to Dr. R. Ham of University of Utrecht for providing *Sedum album* chloroplast DNA clones, to Dr. J. Felsenstein of Washington University for providing PHYLIP, and to Tsukuba Botanical Garden, National Science Museum for providing plant materials.

**References**

ナシ属と近縁な植物における葉緑体 DNA の分化

池 谷 祐 幸
農林水産省果樹試験場
〒305 茨城県つくば市藤本2-1

ナシ属と近縁な数属の植物における葉緑体 DNA の分化について制限酵素断片長多型（RFLP）法によって解析した。その結果、極めて少数の制限サイト多型しか存在しないことが判明し、属間に生する生殖的隔離が弱いことより従来から予測されてきた分類群間に於ける遺伝的分化のレベルは著しく低いであろうという仮説を裏付けた。ナシ属の種間に於ける葉緑体 DNA の分化は形態的な分化とは相関しておらず、浸透交雑などによる葉緑体ゲノムの移入が示唆された。

＜質疑応答＞
松浦（トホク種苗）：バラ科のゲノム分析はどの程度なされていますか。
池谷：バラ科の果樹は普通2倍体で、コムギのようなゲノムの複合による分化は起きていない。
A SIMPLE METHOD TO CLASSIFY RICE STRAINS WITH AA GENOME
AND TO INFER THEIR RELATIONSHIPS BY IDENTIFICATION OF
TRANSPOSABLE ELEMENTS AT VARIOUS LOCI

Eiichi Ohtsubo*, Kayoko Mochizuki, Tyoaki Tenzen and Hisako Ohtsubo

Institute of Molecular and Cellular Biosciences, University of Tokyo,
Yayoi 1-1-1, Bunkyo-ku, Tokyo 113, Japan
* To whom correspondence should be addressed
Fax and Tel. 03-5684-3269

Introduction

Rice belongs to the genus *Oryza* that includes species closely related to the two
cultigens, *Oryza glaberrima* Steud., the cultivated rice of West Africa, and *Oryza sativa*
L., the main rice species cultivated world wide. All taxa of this *O. sativa* complex are
considered to have the same genome, AA, since their F₁ plants have shown no signifi-
cant disturbances in chromosome pairing (Morinaga 1964; Chu et al. 1969). They are
*O. sativa*, *O. nivara*, *O. rufipogon*, *O. glaberrima*, *O. barthii*, *O. longistaminata*, *O.
glumaepatula* and *O. meridionalis* (see Vaughan 1989). In these, the Asian wild rice
species, *O. rufipogon*, is considered to be the ancestor of the cultivated rice, *O. sativa*
(see Oka 1988), and is distributed in tropical Asia from South China to West India and
Indonesia. This species is differentiated into annual and perennial populations (Morishima et al. 1984), and annual type is classified as a distinct species under the name of
*O. nivara* Sharma. Since these ecotypes are interfertile, however, it is considered that
they belong to a single biological species (Morishima et al. 1984).

Morphological and cytological studies and ecological genetic studies of the genus
*Oryza* have helped explain species relationships. Recently, isozyme studies have used
extensive wild species germplasm (Second 1982; 1985). With the current interest in
wild species, application of techniques such as restriction fragment length polymor-
phisms and studies at the DNA level might have enhanced our knowledge of *Oryza*
germplasm.

We present here a simple and useful method to classify and to infer relationships
of rice strains belonging to seven species with AA genome in the *Oryza* genus. The
method is based on the identification of transposable elements at the respective loci in
the rice chromosomes. For the actual experiments, the presence or absence of various
kinds of transposable elements or their members at the respective loci in the rice
chromosomes is determined by using techniques including polymerase chain reaction (PCR). We demonstrate that some elements are present at the corresponding loci in all the strains examined, but the others are not. The patterns for the presence and absence of the transposable elements at the respective loci enabled us to classify the rice strains with AA genome into ten groups and to infer their relationships.

**Transposable elements useful for classification of rice strains**

1) **Retroposons, p-SINE1 and Ret1.** Retroposons are transposable elements that are integrated into a different position via cDNA intermediates (for a review, see Deininger 1989). In plants, an element (designated p-SINE1) with characteristics of a class of retroposons, called SINEs (short interspersed elements) known in the animal system, has been first described: Two members of p-SINE1 (named r1 and r2) are located in two different introns in the Waxy gene in *Oryza sativa*, respectively (Umeda et al. 1991; Fig. 1A), and five members of p-SINE1 (named r3, r4, r5, r6 and r7) are located at the five respective loci in *O. sativa* (Mochizuki et al. 1992) (Fig. 1A). These sequences have a T-rich pyrimidine tract at their defined 3' end and are flanked by direct repeats of a sequence of mostly 14-15 bp long (Fig. 1A). The consensus sequence derived from seven members is 123 bp in length and has an internal promoter region for RNA polymerase III (Fig. 1A). In the seven p-SINE1 members, r2 and r6 are not present at the corresponding loci in *O. glaberrima* (Umeda et al. 1991; Mochizuki et al. 1992).

During the course of the experiment to examine the presence or absence of p-SINE1-r4 in the rice chromosomes by PCR, we have observed that two *O. longistaminata* strains generate the fragments that are slightly larger than those only with p-SINE1-r4 (Mochizuki et al. 1993). Cloning and sequencing of the fragments have revealed that there exists an insertion of a 28-bp sequence near p-SINE1-r4. A 15-bp sequence on one side of the insertion also appeared at the other side, such that the two 15-bp sequences flank the 13-bp sequence (Fig. 1B), indicating that the 13-bp sequence has been transposed to the 15-bp sequence used as target to give rise to duplication of the target sequence. Unlike the other transposable DNA elements to be described below, the 13-bp sequence is not flanked by terminal inverted repeat sequences. There is, however, a T stretch at an end of the 13-bp sequence (Fig. 1B), suggesting that the sequence has been inserted by retroposition. We thus named this sequence Ret1 (for retroposed element #1) (Mochizuki et al. 1993).

2) **Transposons, Tnr1 and Tnr2.** Transposable DNA elements or transposons have terminal inverted repeats (TIRs) and make target sequence duplication (TSD)
Fig. 1. (A) a, Nucleotide sequences of the end regions of p-SINE1 members, r1~r7, and their flanking regions (Mochizuki et al. 1992). The p-SINE1 sequences are indicated by boldface letters. Slashes indicate gaps introduced to improve the alignment. Note that there exist direct repeat sequences at the regions flanking each p-SINE1. b, Nucleotide sequences of p-SINE1 members. A consensus sequence derived from the sequences of all p-SINE1 members is shown by boldface letters at the top. The sequences of seven members are shown under the consensus sequence. Dashes indicate nucleotides identical to those in the consensus sequence. Slashes indicate the bases deleted. The internal RNA polymerase III promoter, A-box and B-box (Galli et al. 1981), is indicated above the consensus sequence. Coordinate numbers (1~123) given to p-SINE1 are shown. (B) The nucleotide sequence of Ret1 and its flanking regions (Mochizuki et al. 1993). The Ret1 sequence (boldface letters) is flanked by direct repeats of a 15-bp target sequence. Numbers (1~13) are coordinates given to Ret1.

upon transposition. Transposons in plants have been classified by homology of their TIRs and the length of TSDs into three families, Ac, CACTA and Mu (for reviews, see Fedoroff 1989; Gierl and Sædler 1992). Each of them is composed of autonomous and non-autonomous elements. An autonomous element is a complete form of the transposon capable of transposing by itself, whereas a non-autonomous element is defective, capable of transposing only with the help of transposase encoded by the autonomous element.

We have recently identified from the Waxy gene in O. glaberrima a 240-bp sequence named Tmr1 (Umeda et al. 1991), whose copy number was estimated to be
about 3500 per haploid genome (TENZEN et al. 1994). Six members of Tnr1 located at different loci in O. glaberrima and O. glumaepatula have been isolated (TENZEN et al. 1994). The Tnr1 members were found to be similar in size and highly homologous (about 80%) to the Tnr1 sequence identified first in the Waxy gene (Fig. 2A). A

![Diagram](https://example.com/diagram.png)

**Fig. 2.** (A) a, Nucleotide sequences of the end regions of Tnr1 members and their flanking regions (TENZEN et al. 1994). The Tnr1 sequences are indicated by boldface letters. Tnr1A, Tnr1E and Tnr1F were identified in O. glaberrima, while Tnr1B, Tnr1C and Tnr1D were identified in O. glumaepatula. Tnr1W, which has been previously identified in the Waxy gene in O. glaberrima (ÜMEDA et al. 1991), is also shown for comparison. Sequences 5'-PuTA-3' and 5'-TApy-3' are commonly seen at the junctions with Tnr1 members. b, Nucleotide sequences of Tnr2 members. A consensus sequence derived from the nucleotide sequences of all the Tnr1 members is shown at the top by boldface letters. Dashes indicate nucleotides identical to those in the consensus sequence. Slashes indicate the bases deleted. Coordinate numbers (1-235) given to Tnr1 are shown. Terminal inverted repeats in Tnr1 are shown by a pair of arrows. (B) Nucleotide sequences of Tnr2 and its flanking regions (MOCHIZUKI et al. 1993). The Tnr2 sequence (boldface letters) is flanked by direct repeats of a 8-bp target sequence. Coordinate numbers (1-157) given to Tnr2 are shown. Terminal inverted repeats in Tnr2 are shown by a pair of arrows.
consensus sequence of Tnr1 (235 bp in length), which is derived from the nucleotide sequences of all the Tnr1 members, has terminal inverted repeat sequences of 75 bp (Fig. 2A). Almost all the chromosomal sequences that flank the Tnr1 members are 5'-PuTA-3' and 5'-TAPy-3' (Fig. 2A), indicating that Tnr1 transposes to 5'-PuTAPy-3' duplicating the TA sequence. Some rice species do not contain the Tnr1 members at the corresponding loci in their chromosomes. Comparison of nucleotide sequences of the regions with or without a Tnr1 member have supported the idea that Tnr1 transposes to 5'-PuTAPy-3' duplicating the TA sequence (TENZEN et al. 1994). Tnr1 is supposed to be a defective form of an autonomous element capable of transposing by itself (TENZEN et al. 1994).

We have also recently found that an insertion of a 157-bp sequence, named Tnr2, is present in the p-SINE1-r4 locus of O. sativa Indica and O. glaberrima but not in O. sativa Japonica (MOCHIZUKI et al. 1992; 1993). The sequence has terminal inverted repeats of about 56 bp long (Fig. 2B) and is present in many copies at various loci only in the rice strains. Tnr2 is present in the 3'-end region of p-SINE1-r4, such that an 8-bp sequence in p-SINE1-r4 appears twice at the junctions, indicating that Tnr2 generates target duplication of 8 bp (MOCHIZUKI et al. 1993). Tnr2 seems to be a defective form of an autonomous element because of its small size. Tnr2 as well as Tnr1 shows no sequence homology with the other transposable DNA elements in plants described previously (FEDOROFF 1989; GIROLI and SAEDLER 1992; BUREAU and WESSLER 1992).

Identification of transposable elements at the respective loci in rice strains with AA genome

We have analyzed the presence or absence of all the seven p-SINE1 members (r1 to r7) located at the different loci in rice strains belonging to seven species with AA genome by PCR using primers that hybridize the sequences flanking each transposable element at the respective locus (see Table 1). The amplified fragments, which were generated from all the strains (see Table 2 for a list of rice strains examined), were unique in size for each of the four p-SINE1 members (r1, r3, r5 and r7) and were hybridized with the p-SINE1 probe (MOCHIZUKI et al. 1993; HIRANO et al. 1994) (see Table 2). This shows that the four p-SINE1 members are present at the corresponding loci in all the strains examined. Similarly, Tnr1W, a Tnr1 member in the Waxy locus, was found to be present in all the rice strains examined (Table 2).

p-SINE1-r2 was present in the Waxy gene of most of the Asian strains including all the O. sativa strains and of a fraction of the O. rufipogon strains (MOCHIZUKI et al. 1993; HIRANO et al. 1994) (see Table 2). Another p-SINE1 member, r6, was not present
Table 1. Synthetic oligonucleotide primers for PCR that are useful to identify the presence or absence of transposable elements at the respective loci in the rice chromosomes

<table>
<thead>
<tr>
<th>Primera</th>
<th>Sequence</th>
<th>Positionb</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>5'-CATGCATCTTTTCATTGTCGT-3'</td>
<td>-182~164</td>
</tr>
<tr>
<td>F2</td>
<td>5'-GGAGGAGCTCAATCGTCC-3'</td>
<td>-213~194</td>
</tr>
<tr>
<td>F3</td>
<td>5'-GCCAAGGTTGAAACTCACC-3'</td>
<td>+173~152</td>
</tr>
<tr>
<td>F4</td>
<td>5'-ATTGCTCCCTACATATGCACTCACC-3'</td>
<td>-103~79</td>
</tr>
<tr>
<td>F5</td>
<td>5'-GACAGTGTTAGAAGGTAT-3'</td>
<td>-57~40</td>
</tr>
<tr>
<td>F6</td>
<td>5'-GTATCTAAGGCGTACCGCTTGACT-3'</td>
<td>-134~110</td>
</tr>
<tr>
<td>F7</td>
<td>5'-GTGATGTGTTACACTGATTAGCT-3'</td>
<td>-94~74</td>
</tr>
<tr>
<td>B1</td>
<td>5'-CTCCTGATCTGTTCTCCATGTC-3'</td>
<td>+177~198</td>
</tr>
<tr>
<td>B2</td>
<td>5'-TGCCCGGATCTCTCTCCATCTC-3'</td>
<td>+191~110</td>
</tr>
<tr>
<td>B3</td>
<td>5'-GTATACATTGAGCTGCACT-3'</td>
<td>+309~327</td>
</tr>
<tr>
<td>B4</td>
<td>5'-AACCATCTAAAGAACAC-3'</td>
<td>+146~152</td>
</tr>
<tr>
<td>B5</td>
<td>5'-GCAGAATGGGATGATTTTGCAGC-3'</td>
<td>+269~293</td>
</tr>
<tr>
<td>B6</td>
<td>5'-AGTAGCAAGAATTTGGGCACT-3'</td>
<td>+259~279</td>
</tr>
<tr>
<td>B7</td>
<td>5'-GAGTTGGATGACTTACATAC-3'</td>
<td>+223~241</td>
</tr>
<tr>
<td>MW81</td>
<td>5'-AGAGGAGCTTCAGAAGACTGCA-3'</td>
<td>4275~4295</td>
</tr>
<tr>
<td>MWR91</td>
<td>5'-AGCACATCTCTCCAGATTCCATTC-3'</td>
<td>4704~4784</td>
</tr>
</tbody>
</table>

* Primers F1 and B1 (or F2 and B2, etc.) hybridize respectively to one strand and its complementary strand of the sequences flanking each p-SINEI member (r1, or r2, etc.). Note here that primers F4 and B4, which hybridize the flanking sequences of p-SINEI-r4, can also be used to identify the presence or absence of Trn2 or Ret1, which is located within or near p-SINEI-r4, respectively (Mochizuki et al. 1993). Primers MW81 and MWR91 are used to amplify the fragment containing Trn7W, a Trn1 member at the Waxy locus (Mochizuki et al. 1993; Tenzen et al. 1994).

b Numbers of positions of primers defined by taking the 5'-end of p-SINE1 as position 1 (see Fig. 1A), as described previously (Mochizuki et al. 1992). Positions of primers MW81 and MWR91 are those of the Waxy gene described by Umeda et al. (1991).

at the corresponding locus in some rice strains in Africa including five strains of *O. glaberrima* and four strains of *O. barthii* (Table 2). In the other African rice strains, one strain (C8595) of *O. glaberrima* and two strains (W0031 and W1650) of *O. longistaminata* generated both fragments with and without p-SINE1-r6, whereas the remaining two *O. barthii* strains and five *O. longistaminata* strains generated only the fragments with p-SINE1-r6 (Table 2).

The last p-SINE1 member, r4, was present at the corresponding locus in all the strains examined (Table 2), but the amplified fragments showed a length polymorphism due to the presence or absence of Trn2 or Ret1, as mentioned earlier. Trn2 was present within the sequence of p-SINE1-r4 in the restricted strains, in which three *O. longistaminata* strains and one *O. rufipogon* strain generated the fragments without Trn2 in addition to those with Trn2 (Table 2). Ret1 was located near p-SINE1-r4 and was present only in two *O. longistaminata* strains, W1443P2 and W1650, in which the latter had the p-SINE1-r4 sequence without Ret1 in addition to that with Ret1 (Table
### Table 2. Presence of transposable elements in various *Oryza* strains with AA genome.

<table>
<thead>
<tr>
<th><em>Oryza</em> Species</th>
<th>Strain (Type)</th>
<th>Habit</th>
<th>Origin</th>
<th>p=SINE1</th>
<th>Tur/W</th>
<th>Tur3</th>
<th>Ret1</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>sativa</em></td>
<td>T65 (JA)</td>
<td>Taiwan</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Nipponbare (JA)</td>
<td>Japan</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Susumihiko (JA)</td>
<td>Japan</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>S48 (JA)</td>
<td>Japan</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>552 (JA)</td>
<td>Japan</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>709 (JA)</td>
<td>China</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>1021 (JA)</td>
<td>China</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>306 (JV)</td>
<td>Indonesia</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>538 (JV)</td>
<td>Japan</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>654 (JV)</td>
<td>Indonesia</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>718 (JV)</td>
<td>China</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>444 (IN)</td>
<td>India</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>1 (IN)</td>
<td>Vietnam</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>108 (IN)</td>
<td>Taiwan</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>C340 (IN)</td>
<td>India</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>423 (IN)</td>
<td>India</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>721 (IN)</td>
<td>China</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>B</td>
</tr>
<tr>
<td><em>rufipogon</em></td>
<td>W030</td>
<td>Myanmar (Mawla)</td>
<td>+</td>
<td>+</td>
<td>+ +</td>
<td>+</td>
<td>-</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>W1536</td>
<td>Sri Lanka (Putalum)</td>
<td>+</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1981</td>
<td>Sri Lanka</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W1997</td>
<td>India, west</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W2010</td>
<td>India</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W140</td>
<td>India (Raipur)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1204</td>
<td>Pakistan (Kanpur)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1906</td>
<td>Sri Lanka</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1970a</td>
<td>Indonesia (Java)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1649</td>
<td>Thailand (Bangkok)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W0252</td>
<td>India (Orissa)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W0133</td>
<td>India (Birarol)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W0136</td>
<td>India (Kodim)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W0172</td>
<td>Thailand (Chumphon)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1992A</td>
<td>Thailand, south</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1843</td>
<td>China, central</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1956</td>
<td>China</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W2053</td>
<td>Indonesia (Java)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>X1</td>
<td>China, south</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td><em>glaberrina</em></td>
<td>GMS1 [E106]</td>
<td>[Bred line]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W410</td>
<td>French Guinea</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>C7599</td>
<td>Nigeria (Badeggi)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>C8257</td>
<td>Sierra Leone (Kambia)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>C8389</td>
<td>Guinea (Kinda)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>C8595</td>
<td>Cameroon (Yapoua)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td><em>barthii</em></td>
<td>W0822</td>
<td>Mali (Kogon)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W0062</td>
<td>Ghana</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1416</td>
<td>Sierra Leone (Kaeowa)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>+</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>W1467</td>
<td>Cameroon (Yapoua)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1581</td>
<td>Chad (N'Djamena)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W0653</td>
<td>Sierra Leone (Konta)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1438</td>
<td>Chad (N'Djamena)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1426</td>
<td>Mali (Maqit)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1451P1</td>
<td>Ivory Coast (Korhogo)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1415P2</td>
<td>Nigeria (Soloto)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1443P2</td>
<td>Mali (Tomboctou)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1650</td>
<td>Tanzania (Zanzibar)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W0631</td>
<td>Sudan</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1414P1</td>
<td>Sierra Leone (Monu)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td><em>glaucescens</em></td>
<td>W1169</td>
<td>Cuba (Santa Clara)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td><em>meridionalis</em></td>
<td>W1417</td>
<td>Brazil (Belen)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1237</td>
<td>Australia (Darwin)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1636</td>
<td>Australia (Darwin)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>W1625</td>
<td>Australia (Darwin)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

*a* The presence (+) or the absence (−) of each transposable element in rice strains was determined by the analyses of the PCR-amplified fragments with or without it, respectively, by agarose gel electrophoresis, followed by Southern hybridization with appropriate probes and/or by DNA sequencing (UMEDA et al. 1991; MUCHELI et al. 1993; HIRANO et al. 1994). ± indicates the strain that generated both fragments with and without a member of a transposable element. Primers used to amplify the fragments by PCR are listed in Table 1.

*b* Strains listed are in the collection at the National Institute of Genetics, JA, Japonica; IV, Javanica; IN, Indica.

*c* P, perennial type; A, annual type; I, intermediate or weedy type.
Note here that Tnr2 was present in the corresponding locus in five of the six *O. sativa* Indica strains but not in all of the *O. sativa* Japonica and Javanica strains examined (Table 2). *O. sativa* Japonica is more closely related to *O. sativa* Javanica than to *O. sativa* Indica (GLASZMANN 1987). Tnr2 inserted in p-SINE1-r4 must, therefore, be a very useful character classifying *O. sativa* Indica from *O. sativa* Japonica or Javanica. Note also that RetI is useful to identify and distinguish some strains belonging to *O. longistaminata*.

**Classification of rice strains with AA genome into ten groups by examining the presence of transposable elements at respective loci in the rice chromosomes: Classification identifies wild rice species most related to the cultivated rice species.**

As described above, the p-SINE1 members at five loci (r1, r3, r4, r5 and r7) and a TnrI member are found in all of the rice strains with AA genome, whereas two p-SINE1 members (r2 and r6), Tnr2 and RetI, are in the restricted strains. These findings would confirm that p-SINE1-r2, p-SINE1-r6, RetI and Tnr2 had been inserted to each locus during divergence of the species with AA genome, but the others had been inserted to the respective loci before the divergence. From the results shown in Table 2, the rice strains examined were classified into 10 groups (A–J) (MOCHIZUKI et al. 1993) (see Table 3). In this classification, the strains belonging to one species are

**Table 3. Classification and relationships of rice strains with AA genome.**

<table>
<thead>
<tr>
<th>Groupa</th>
<th></th>
<th>p-SINE1</th>
<th></th>
<th></th>
<th>Oryza species (Number of strainsb)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r1, 3, 4, &amp; 5 and 7</td>
<td>r2</td>
<td>r6</td>
<td>Tnr2</td>
<td>RetI</td>
</tr>
<tr>
<td>A</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>B</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>D</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>E</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>F</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>G</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>H</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>I</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>J</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

a Each group shows a unique pattern for the presence or absence (+ or −, respectively) or both (±) of transposable elements. Groups were placed side by side such that one group is most similar to the next group (MOCHIZUKI et al. 1993). Note that group A is next to group J.

b See Table 2 for the rice strains examined.
not always in the same group: The *O. sativa* strains are divided into two groups (A and B), the *O. barthii* strains into two groups (C and D), the *O. rufipogon* strains into four groups (A, B, F and J), and the *O. longistaminata* strains into five groups (F–J) (Table 3). The strains in a group often belong to two or more species (Table 3).

It seems that each group contains the strains belonging to the species which are closely related each other: Group A contains two strains of *O. rufipogon* and all the strains of *O. sativa* Japonica and Javanica, and exceptionally one strain (444) of *O. sativa* Indica (Table 3), which has been shown to have an intermediate isozyme polymorphism pattern between those of the *O. sativa* Indica and Japonica strains (GLASZMANN 1987): Group B contains the strains of *O. rufipogon* and *O. sativa* Indica: Group D contains the strains of the two species in Africa, *O. glaberrima* and *O. barthii*, which are considered to be closely related each other (CHU et al. 1969). These lead us to consider that the rice strains in one group are very closely related to one another in general, even though they have been previously classified to be in different species. According to this consideration, the *O. glumaeapatula* strains in America, the *O. meridionalis* strains in Oceania and some *O. barthii* strains in Africa, all of which are classified into the same group C (Table 3), are very closely related to one another. The other *O. barthii* strains belonging to group D must be distinct from those in group C (Table 3). Three *O. longistaminata* strains in Africa and one *O. rufipogon* strain in India are in the same group F, whereas one *O. longistaminata* strain and four *O. rufipogon* strains (Table 3), which are distributed in South Asia and China, but not in India, are in the same group J (see Table 2), indicating that the strains in group F or J are very closely related to one another.

The *O. rufipogon* strains that are divided into four groups, A, B, F and J, of which only groups A and B include all the Indica and Japonica (Javanica) strains of *O. sativa*. It is interesting that *O. rufipogon* which is the most probable ancestral species of cultivated rice, *O. sativa* (OKA 1988), seems to have been differentiated in a corresponding manner to the *O. sativa* subdivision. This leads us to speculate that the *O. sativa* strains in groups A and B have been derived respectively from the *O. rufipogon* strains in the corresponding groups. Noteworthy is that the *O. rufipogon* strains in group A are distributed in Thailand, while those in group B that include both the annual and perennial types are distributed in India and South Asia, but not in China (see Table 2).

**Relationships among rice strains with AA genome.**

The groups shown in Table 3 are actually arranged side by side by their patterns for the presence and absence of transposable elements: The pattern of group A is most similar to the pattern of either one of the adjacent groups, B and J; the pattern of group
B is most similar to the pattern of either one of the adjacent groups, A and C; and so on.

It seems that the strains in the two groups placed side by side are more closely related than those in the other groups. Indeed, the O. sativa Japonica and Javanica strains in group A are most closely related to the O. sativa Indica strains in group B, and the African cultivated rice strain(s) of O. glaberrima in group D are more closely related to the African wild rice strains of O. barthii in group C than to the Asian cultivated rice strains of O. sativa in groups A and B. These lead us to consider that the rice strains in the two groups placed side by side are more closely related than those in the distant groups in general. According to this consideration, the O. rufipogon strains in group F or J are closely related to the O. longistaminata strains. The strains of O. glumaepatula, O. meridionalis, and O. barthii in group C are related more closely to O. sativa Indica than to O. sativa Japonica (Javanica).

We have observed that an O. glaberrima strain (C8595) in group E generates both fragments with or without p-SINE1-r6 (Tables 2 and 3). This strain has long awn, a character of O. barthii (OKA et al. 1978), and was found in a population where O. glaberrima and O. barthii coexisted (Y. SANO unpublished result). Since O. glaberrima, a cultivated species, and O. barthii, its wild relative, produce fertile hybrids (CHU et al. 1969), it is likely that C8575 is a hybrid between an O. glaberrima strain without p-SINE1-r6 and an O. barthii strain with p-SINE1-r6. Likewise, five of the seven O. longistaminata strains generate the fragments with or without p-SINE1-r6, Tnr2 or Ret1 (Tables 2 and 3). These O. longistaminata strains that grow vegetatively must have carried the elements in the heterozygous state.

**Final notes**

There have been no reports that the SINE elements are excised from the loci, where they have been inserted. We have observed that five of the seven p-SINE1 members are present in the corresponding loci in all the rice strains examined, supporting further that p-SINE1s are not excised. It has been reported, however, that the transposable DNA elements can be excised from their original sites by the action of transposase encoded by the autonomous elements (NEVERS et al. 1986). The excision events are usually imprecise and give different sequences from the original target sequences by deletion or substitution of a DNA segment. We have previously shown that the DNA segments without Tnr2 do not contain deletions or substitutions at the target sequence (MOCHIZUKI et al. 1993), indicating that Tnr2 has not been excised often during divergence of the rice strains with AA genome. As described above, TnrIW, a Tnr1 member at the Waxy locus, is present in all the rice strains examined.
(Table 2), indicating that Tnr1 is also not excised often during divergence of the rice strains.

We believe that if any one of transposable elements including those described here could be identified at many other loci in the rice strains with AA genome, they could become useful characters not only for classifying the rice strains but also for inferring their relationships in more details.

Acknowledgments

This work was supported by a grant Pioneering Research Project in Biotechnology from the Ministry of Agriculture, Forestry and Fishery of Japan.

Summary

We present here a simple and useful method to classify rice strains with AA genome and to infer their relationships. The method is based on the identification of transposable elements at various loci in the rice chromosomes by considering that members of the transposable elements must have been inserted into the respective loci during divergence of the rice species. Here we first describe several kinds of transposable elements identified in our lab that are very useful for this study. These are retroposons, including a SINE element (called p-SINE1) and Ret1, and two transposons, Tnr1 and Tnr2. We then show our recent results of analyses of the presence or absence of these elements in various rice strains belonging to seven species with AA genome in the Oryza genus. Five p-SINE1 members (r1, r3, r4, r5 and r7) and a Tnr1 member were present at the corresponding loci in all the strains examined, but two p-SINE1 members (r2 and r6), Ret1 and Tnr2 were not present at the corresponding loci in some strains. From the results obtained, the rice strains examined could be classified into 10 groups (A–J). The strains belonging to one species are not always in the same group, and the strains in a group often belong to two or more species. Each group, however, often contained the rice strains belonging to the species which have been assumed to be closely related each other. When the groups are arranged side by side by their patterns for the presence and the absence of transposable elements, such that the pattern of group A is most similar to the pattern of either one of the adjacent groups, B and J, the strains in the two groups placed side by side seemed to be more closely related than those in the distant groups.
References


＜質疑応答＞
田中（生物研）：Tnr のイネゲノム中に多数のコピーがあるそうですが、実際のイネの栽培品種は自然突然変異をほとんど起こしません。これはこれらのトランスポゾンがほとんど転移しないというふうに考えてよろしいのですか。
大坪：そう考えてよい。トランスポゾンで入ると安定な mutant となる。有名なエンドウのシワやキンギョソウの白花はトランスポゾンによるものであることが分かった。
丹羽（茨城大）：SINE1 の存在はモチーウルチとは関係ないと考えて良いか。
大坪：はい。関係ないと考えられます。
SEXUALITY AND ACCUMULATION OF DELETERIOUS MUTATIONS IN HUMANS

Naoyuki TAKAHATA

Department of Genetics, The Graduate University for Advanced Studies,
Mishima 411, Japan

I first discuss the evolutionary significance of sexuality. My most arguments follow J. F. Crow who has recently provided a concise review of this rather old issue that was first raised by H. J. Muller in the 1930s. Here it goes.

A species that reproduces asexually, such as a dandelion, can respond rapidly to selection, for it has a very high heritability. Except for environmental influences, the progeny are exactly like their parents. But the other side of the coin is that the future amount of change in the population is limited by the phenotype of the most extreme individual in the population. Usually mutation is too slow to make up for the decreased variability. Thus the effectiveness of selection in an asexual species is limited.

A sexual species, with its free Mendelian segregation and recombination, has no such limitation. Recombination among the genes within a population can generate types that are far beyond the range of any individual in the current population. Thus a change in environment may lead to very large changes. It is the ability to scramble genes and try new combinations that gives sexual reproduction such an evolutionary advantage. This is borne out by the paleontological data. Many asexual species are highly successful (like dandelions). But asexual species are relatively short-lived when long geological periods are considered, and contemporary asexual species had sexual ancestors.

In sexual species mutation plays an essential role in providing the ultimate source of genetic variability, but the rate of evolution is not tightly coupled to the rate of mutation in the evolution of phenotypes. All that seems to be required is a bare minimum. Selection and recombination play a more decisive quantitative role. If mutation were to stop entirely, the population would retain variability and be able to respond to selection for many generations. Unless selection were very intense, this could be hundreds or even thousands of generations. Advantages of sexual reproduction become clearer when environment changes.

However, the great majority of mutations that affect function in any way are
harmful. They are eliminated by natural selection rather quickly in terms of evolutionary time, and make no permanent contribution to the evolution of the species. They do, however, affect the fitness of the individuals of the population at any particular time. A harmful mutation persists in the population until all the descendants of the original mutation are eliminated by natural selection, or chance (random genetic drift). How long the mutation persists depends on how harmful it is and how it interacts with other genes, especially on whether it is dominant or recessive. A harmful dominant mutation is eliminated rather quickly. If it reduces the viability or fertility of its carrier by 10 percent it will persist in the population on the average for 10 generations. In contrast, a recessive mutation may persist for hundreds or thousands of generations before there are enough homozygous descendants to eliminate the mutation from the population. Actually, however, completely recessive genes are rare. Almost always the gene, even if seemingly recessive, has some effect in the heterozygous state. Usually it is these heterozygous effects that cause the gene to be eliminated from the population. The reason is that homozygotes are so rare that even a small selective disadvantage of the much more common heterozygotes has a bigger net effect. In *Drosophila*, the only species where the answer to this question is known, the average "recessive" mutation has enough heterozygous expression to persist in the population about 50 generations. Sexual reproduction is efficient to eliminate deleterious mutations as clusters for it allows to produce extreme individuals who carry them by a combinatorial means. On the other hand, asexual reproduction does not for offspring always carry deleterious mutations no smaller than their parents. There is a steady increase of deleterious mutations in the genome, and this steady deterioration of the genome under a constant environment was what Muller argued. Hence the presumed phenomenon is called the Muller's ratchet.

Under changing environment, deleterious mutations do not always deleterious. For instance, the relaxation of natural selection in recent times is conspicuous particularly in countries with advanced technology. It means almost certainly that deleterious mutations are accumulating faster than they are being eliminated. We don't notice this by such things as increased illness and death rate because the genetic weaknesses are compensated by improved environment. At the moment environmental changes are in the lead, for health, survival, and life expectancy continue to increase. The question is: Can we keep improving the environment enough to stay ahead of the deleterious effects of an increased accumulation of mutations. The answer is not obvious. But one thing is clear: If we for some reason had to go back and live the way our ancestors did 10,000 years ago we would have a miserable time. We would suffer all the ills that they did, plus the weakening effects of mutations that have accumulated in the intervening years as selection was relaxed.
It is very likely that accumulated deleterious mutations will eventually be a serious problem for mankind. Fortunately any such change will be slow. If we can judge from *Drosophila* the changes from our present high standard of living and relaxed selection will gradually manifest themselves over a period of some 50 generations. This is about 1500 years, perhaps time enough for humanity to learn more efficient ways of coping with the problem—provided we are not overwhelmed by a runaway population increase.

Now I present my view not only that the relaxed selection and the accumulation of deleterious mutations have been occurring in our days, but also that it has actually begun some million years ago in the lineage leading to modern humans.

Available genetic data reveals that the human population is more variable than the chimpanzee population at the protein level, whereas limited data indicates that the opposite is the case at the DNA level. The lower level of silent polymorphism in the human population suggests that its long-term effective size is smaller than the chimpanzee’s. The past 25 years or so have witnessed a wide application of electrophoresis to various problems in evolutionary biology. A puzzling finding was that the extent of genetic variation in humans is highest among primates, or even higher than that in most vertebrates. The heterozygosity (H) is 14.3% when averaged over 121 loci of humans. The value is much higher than that of non-human primates, which is 1.3% over 43 loci of chimpanzees and 4.6% over 22 loci of gorillas, and as high as that of some *Drosophila* species whose $N_e$ is said to be 100 times as large as that of humans. These figures may occur simply by chance if different loci are used in different organisms, because the degree of polymorphism at the amino acid level differs greatly from locus to locus. Another possible cause is different mutation rates among different taxonomic groups. The mutation rate may well depend on reproductive ages and/or physiological conditions. To avoid these uncertainties, it is sensible to compare H for the same set of loci among relatively closely related primates. Chimpanzees, the closest relative to humans, are clearly the best organism to be compared in the present context. There are 35 loci that were examined commonly to both species with reasonably large sample sizes. It turns out that the average H value is 8.0% for humans and 2.3% for chimpanzees. There is still a three- to four-fold difference between these H values.

It is interesting to evaluate the fraction of selectively neutral mutations, $f_\alpha$, which contribute to H. One method uses rare variant alleles. These are alleles whose frequencies are smaller than a pre-assigned value $q$, usually taken as 0.01 or less. Because of their rarity, rare alleles are less likely to be subject to natural selection. The average sample size of individuals for the above 35 loci is 111. This sample size is relatively small and the value of $q$ chosen as 0.05 is relatively large. Nonetheless,
these limitations should not alter the conclusion, because the number of rare variant alleles of \( q \leq 0.01 \) is also much larger in humans than in chimpanzees (11 vs 3). The estimated \( f_o \) is 43% for humans and 21% for chimpanzees. The rare-allele method was previously applied to much larger data sets for two ethnic groups, giving that \( f_o=14\% \) for European and 21% for American. Both estimates are smaller than the present estimate of 43%, but this may be again due to different loci examined. In any event, all this indicates that the fraction of selectively neutral mutations is larger in humans than in chimpanzees and that the remaining 57% and 79% of mutations are definitely deleterious in humans and chimpanzees, respectively. Hence, in terms of both \( H \) and \( f_o \), the human population is more variable than the chimpanzee population.

The DNA sequences of alleles at 49 loci in humans were compared. The mean \( \pi \) is 0.04% in the coding region and 0.11% in the noncoding region. In the entire coding region, there are 15 nonsynonymous (amino acid replacement) changes and 11 synonymous changes, corresponding to \( \pi_s=0.02\% \) per nonsynonymous site and \( \pi_s=0.08\% \) per synonymous site where subscripts \( S \) and \( N \) stand for synonymous and nonsynonymous sites or substitutions. The ratio of \( \pi_n/\pi_s \), denoted by \( f_n \), is a measure of the average degree of neutrality for nonsynonymous changes (subscript \( n \) stands for the nucleotide level). The estimated \( f_n \) is 31%, indicating that 71% of nonsynonymous changes at these 49 loci are selected against, relative to synonymous changes. The \( f_n \) value is smaller than \( f_o=43\% \), but larger than 14% and 21% for the ethnic groups.

There is a striking difference in the value of \( f_n \) between the human and \textit{Drosophila} populations. Except for humans, \textit{Drosophila} is one of a few genera whose genomes have been extensively examined for the DNA polymorphism. Eleven sequences of \textit{Alcohol dehydrogenase} genes in \textit{D. melanogaster} show that there is only one nonsynonymous change, but the extent of synonymous diversity (\( \pi_s=0.6\% \)) is much larger than that of humans. A large scale survey of DNA polymorphism shows that the \( \pi_n \) (\( \pi \) when measured by restriction enzyme analysis) around 20 loci in the \textit{D. melanogaster} genome ranges from 0% to 0.9%, the average being 0.34%. A general feature is that \( \pi_n \) and \( f_n \) in \textit{Drosophila} are fairly low.

Unfortunately, there are no such comparable data sets for chimpanzees and any other non-human primates. Exceptions are \textit{ABO} \textit{glycosyltransferases} and mtDNA. Partial DNA sequences 405 bp long are available for human \( A \) and \( B \) alleles and there are four nonsynonymous and one synonymous differences. Three chimpanzee \( A \) and two gorilla \( B \) alleles are sequenced. Because the \( B \) antigen is absent in chimpanzees and the \( A \) antigen is absent in gorillas, we may compare chimpanzee \( A \) and gorilla \( B \) alleles. There are four synonymous and four nonsynonymous differences. The ratio of the synonymous to nonsynonymous sites in the sequence is approximately 0.5, so that we obtain \( f_n=0.5 \) for the non-human primates. Although information is limited
and it is likely that some of these nonsynonymous changes are subject to positive selection, the large \( f_n \) value between human \( A \) and \( B \) allele is in accord with \( f_e \) of electrophoretic data. Furthermore, restriction enzyme analysis of primate mtDNA shows that \( \pi_e = 0.4\% \) in humans, 1.3\% in chimpanzees, 0.6\% in gorillas. These \( \pi_e \) values are largely due to silent changes. For the noncoding region of 1135 bp length, several DNA sequences are available to estimate \( \pi_s \); 1.7\% averaged over 14 human mtDNAs and 7.8\% averaged over three chimpanzee mtDNAs. Because this region is known to have evolved much faster than the coding region, the large \( \pi_s \) value is expected. The point is that both restriction enzyme and sequence analyses of mtDNA reveal consistently that the extent of silent polymorphism of mtDNA in humans is one-fourth to one-third of that in chimpanzees. We may note that the relevant mutation rate for discussing \( \pi \) must be measured in units of generations, rather than per year. If the mutation rate per generation has become high in the human lineage because of the prolonged generation time since the evolution of \( H. \) erectus, the above conclusion is reinforced.

Because most, if not all, of silent polymorphism is free from natural selection, the neutral theory suggests that \( N_e \) of modern humans has been one-fourth to one-third of that of chimpanzees for the period during which the currently observed mtDNA variations were generated. The time period concerned with the mtDNA is 0.15 million years (corresponding to the Late Pleistocene), if the female effective number and the generation time have been 5000 and 15 years, respectively. On the other hand, for an autosomal locus, the average persistence time of neutral polymorphism becomes 0.6 million years so that the time period concerned is the Middle Pleistocene epoch.

The original and present form of the neutral theory of M. Kimura assumes that there are only two major classes of mutations; neutral and definitely deleterious. Deleterious mutations do not contribute at all to polymorphism and molecular evolution. In this view, changes in population size are irrelevant to the issue, and changes in selection pressure (the degree of selective constraint, or \( f \)) are the sole possibility to account for the aforementioned molecular data. The neutral theory does not necessarily assume that the degree of selective constraint remains constant throughout a long evolutionary time. As pointed out by Crow, the relaxation in recent times is conspicuous particularly in countries with advanced technology. It means that deleterious mutations are accumulating faster than they are being eliminated, because these mutations no longer are deleterious under the improved environment.

A question is whethere a similar relaxation that can account for the protein polymorphism is supported by paleo-anthropological data. It is known that the human lineage has experienced dramatic changes in brain size, economic and social complexity, and culture over the past 2 million years. Although neither learning nor the mother
-infant relationship is specific to humans, the period of *H. habilis* was an incomparable stage of the development of human characters that was fostered by the enlarged brain size. It meant that infants were born at an earlier stage of mental and physiological maturity. Like brain and the digestive system, the immune system, too, that protects our body from pathogens is not yet fully developed in the newborn baby. Extensive parental care at home bases was needed to nurture such infants and the period of parental care became considerably longer. Associated with the enlarged brain size were naturally anatomical and social changes. Furthermore, during the Pleistocene, repeated glaciations of the Northern Hemisphere occurred and tropical forests were replaced by savanna woodlands. It was essential to solve dietary problems in the face of changing environmental conditions. This might in turn help to develop intelligence, as recently epitomized by C. MILTON as *we are what we ate*. It is conceivable that changed diet together with innovation of tool and fire have altered the selection pressure as well, particularly for the enzymatic system. It therefore seems reasonable to postulate that the enlarged brain had altered the internal and external environments of *H. habilis* and *H. erectus*. Previously deleterious mutations have become neutral or vice versa. However, modern humans appear to be so successful to overcome or improve the environment and therefore it is not absurd to assume that natural selection in effect has been relaxed. In all respects, we are constrained by what our ancestor was challenged to survive. The protein polymorphism, too, could not escape from this constraint.

Genetic data suggest relaxed natural selection in the modern human lineage. An alternative view is provided by the slightly deleterious mutation hypothesis of T. OHTA. However, it has a serious difficulty. Figure 1 illustrates the proportion of slightly deleterious mutations which is required for the hypothesis to be valid. Note also that there is no such a class of mutations in the neutral theory. The slightly deleterious mutation hypothesis inevitably leads to continual deterioration of the human genome, unless the optimum fitness of alleles changes in a short time scale less than a few million years. This genome deterioration is due to the accumulation of slightly deleterious mutations over millions years, and the situation is quite different from that in endangered species such as the cheetah. The reduced fitness of cheetahs results from recessively deleterious mutations unmasked by a sudden decrease in the population size and intense inbreeding. Needless to say, the relaxation hypothesis does not imply that the human genome has been deteriorating. This is true as long as we live under the improved environment. However, it is also true that if we for some reason had to go back and live the way our ancestor (e. g., *Australopithecus*) did, we would suffer the weakening effects of mutations that have accumulated in the intervening years as selection was relaxing. Further characterization of the human genome
necessitates knowledge of the non-human primate genome.

Fig. 1. Relative frequency of three classes of mutations; (N) neutral, (S) slightly deleterious, and (D) deleterious mutations. Those under the neutral theory and the slightly deleterious hypothesis are depicted in upper and lower part, respectively. The frequency in human is represented by open bars and that in chimpanzee by solid bars. In the neutral theory, some 20% of deleterious mutations (D) are converted as neutral and there is no intermediate class of mutations (S). The slightly deleterious hypothesis assumes that whereas the fraction D is the same in both species, the intermediate class (S) is effectively neutral in the human population with a smaller size and is driven by genetic drift. Since this hypothesis does not consider relaxed natural selection, the accumulation of slightly deleterious mutations gradually deteriorates the human genome.
References


＜質疑応答＞
西尾（政育）：動物の MHC と植物の自家不和性に関与する遺伝子の共通性と差異について、また、自己・非自己の認識現象は昆虫や他の生物でも見られるかどうかについて伺いたい。
高畑：多型を積極的に保持するという点は似ているが、分子的には似ていない。他の生物でみられる現象として、ホヤで、違ったコロニーをつけると離れてしまうが、同一系では一つになってしまうという現象がある。
総合討論

座長（鶴田）：マラーによって人為突然変異が発見されたときに2つのことが可能になると考えられた。1つは進化が分かるということ、もう1つは品種改良に使えるということであった。後者については突然変異育種ということで進歩してきて、千数百の品種が世界で育成されてきたが、もう一つの進化ということを人為突然変異によって科学的に解析できるということは、それほど当たりなかったと言える。最近の分子生物学の進歩によって、DNAレベルの進化はだいぶ分かってきたが、人為突然変異と進化との関係はあまりよく分かっていない。今日の御講演で進化についての最近の解明の一端が示された。講演の順に討論をしていきたい。

武田（東京大）：Wx遺伝子は同じ日本型の品種でも違うのが同じなのか。
平野：構造遺伝子はほとんど変わらない。glaberrimaと比較しても1アミノ酸しか違わない。しかし、Wxタンパク質の量は違、アミロースの量が異なる。優良品種とそうでないものでWxタンパク質がどれだけ違うのかは分からない。アミロース含量は米の品質を決める重要な要因であり、それだけでよいかどうかは分からないが、生産性が高い、あるいは耐病性が強いものをアミロース含量を変えることによって、改良することが可能と考える。

長谷川（大阪府大）：Wx遺伝子にみられたイントロンの消失についてもう少し詳しく説明して欲しい。硝酸還元酵素にみられるイントロンについてみると、ある箇所のイントロンは進化の過程で消失しているし、あるイントロンの有無は種間で異なる。特定のイントロンの有無に、遺伝子がコードするタンパク質の構造や機能に関係することもあると思うが、どうだろうか。
平野：イントロンの中にエンハンサー様の活性があれば遺伝子発現に関与するが、どのくらいの機能があるかは一概には言えない。ある遺伝子が進化している過程でイントロンが抜けていったという仮説がある。イントロンがなぜ出来たかということについて、これは真核生物の遺伝子が形成されたときの名残であると考えられている。もともとあるペプチドをコードしている部位がばらばらとなっており、もともとそういう形で遺伝子が作られてきたが、原核生物ではそういうものがすべて抜けていったが、真核生物では抜けていないという説がある。真核生物でも抜けていったものもあるし残っているものもある。この分野で代表的な研究にグロビン遺伝子での研究がある。九州大学の岡先生が、タンパク質のモジュール構造とエクソンが対応していることを見いだされた。そこらではグロビン遺伝子では2つのイントロンがあることが分かっていたが、3番目のイントロンを持っているものがずいぶんと見出された。その直後にレグヘモグロビンで3番目のものを見つけられた。その後ギルバートが、多数のイントロンがあるトリオースフォースフェトイソメラーゼの研究を行い、非常に遠くの種で比較し、多くの種でイントロンが11あるが、蚊で1つmissing
イントロンを見いだした。もともとイントロンがあって抜けて行ったと考えられている。一方、イントロンが入ることでは、mobile element が入ったという考えがある。クロピシンの 3 番目のイントロンが生物によって位置が違うので、抜けたと言うより入ったと考えた方がよいというデータも得られている。

生井：植物界一般では、細胞質のオルガネラは母性遺伝が主ですが、お話しの針葉樹では父性遺伝が一般的であるとのこと、大変興味深く勉強させて頂きました。そこで、それらの父性遺伝の機構について分かっていることをお聞かせください。

津村：1970年代にスギで夏に白くなるという色葉の表現型が父方からくることが分かった。その後、葉緑体 DNA を分析すると、すべて父性遺伝することが分かった。筑波大の隅先生により、電顕レベルの研究で、花粉管の中の葉緑体が受精のときに一緒に入り、母方の方が消化されることが示された。なぜそうなるのかは不明。

座長：母性遺伝の場合と父性遺伝の場合はどちら DNA を調べても集団内の GST が変わるということになると思いますが実際にはどうか。

津村：GST は、最初に示した方はアイソザイムで estimate したものですので、両親から来たものをそのまま調べていて、そのあとののは葉緑体ですので、針葉樹については父方だけの系統になってしまうので問題はあるのですが、今使われる情報は少ないので、父方だけとは分かっていながらやっている。

松元（野菜茶試）：池谷さんの研究法で、農業的に有用な形質に着目した、品種・系統レベルでの分類は可能か。また、RFLP 技術を用いる育種的意義はどのようなことがあるか。

池谷：ミトコンドリアでは種内変異があり、かなり識別することが可能である。フィンガープリントのような非常に変わりうるものの個体レベルまで可能であるが、その中間ぐらいのものは、育種的にどう応用できるかは分からない。

松浦（トウホク種苗）：形態的にかなり違うものでも、例えば 2 条オオムギ・6 条オオムギの分類のような場合は、1 遗伝子の差であったが、このようにわゆる分類形質の遺伝子情報はどの程度分かっているのか。また、それぞれについてどの様な重み付けをしているのか。

池谷：種の大きさと遺伝的な差異の程度に関係はない。栽培植物の場合、種の定義が細かくなりすぎている。もっと整理する必要がある。過去の、2 条オオムギ・6 条オオムギの分類や、モモの毛のあるものないものの分類などは問題があったと考えられる。

丹羽（茨城大）：高畑先生にお伺いしますが、自家受精と他家受精を比較した場合、有害遺伝子の蓄積の程度はどうちらが多いか。

高畑：自家受精は無性生殖に近い。つまり、有害遺伝子を蓄積しやすいということになる。

藤本（岐阜大）：被子植物の場合、他家受精植物に変異が保存されていて、自家受精のもののは早く有害変異がなくなっている。他家受精の方が有害変異をハテロに持っており、条件が変わる中立になりまた有利になる。そういう形で変わってきていると考えられるが、これについてご意見をお聞きしたい。

高畑：有害遺伝子といった場合、優劣性が重要で、シミュレーションでお見せしたものは、
semi-dominance のもので、劣性有害という場合は全く違う。多くの有害遺伝子は劣性であるので、ヘテロで発現しているものが非常に多い。どういう遺伝子を論議しているかによって答えは違ってくる。outbreeding している方が変異が多いということについても、自家受精でもいろんなものがいろんな地域で inbreeding している場合は変異が大きい。どういう条件でということを言わないと、outbreeding だから変異は大きいとか、inbreeding だから変異は小さいという様なことはない。

座長：栄養繁殖と有性生殖の場合を比べると、減数分裂を通すことによって、gamete レベルで耐えられないような障害を、栄養繁殖ではそれらを通らなくてもすむということで、有害変異を蓄積するというような効果があると思うかどうか。

高畑：言われるとおりだと思う。gametogenesis を起こすこと自体が、有害遺伝子を除く効果を持っている。

久保山（東京大）：イネの系統間で SINE のコピー数は相当異なるものか。また、話のあったトランススポゾンは一定の率で発現するものか、ある条件で動きやすくなるものか。

大坪：glaberrima と sativa の正確なコピー数はまだ測定していない。これらは sativa から取ってきたものであるが、glaberrima にあって sativa にないものはまだ見つかっていない。sativa の方が多いと考えられる。longistaminata もほとんど glaberrima と同じ。2つ目の質問で、いつ跳んだかについては、各種で独立に跳んでいると思っていたが、実際には、ある跳ぶ時期があってそのときに広がったのではないかと考えられるのがよく分からない。もっと locus を増やして調べる必要がある。

座長：トランススポゾンの起源はどういうふうにお考えですか。

大坪：レトロトランススポゾンの場合はレトロウィルスから出てきたと考えられるが、その逆かも知れない。他のエレメントは構築的な共通性はあるが塩基配列は全く違う。生物進化の過程でそれぞれ平行的に生じたと考えられるが、水平移動もあると考えている。