Review

Gene content, organization and molecular evolution of plant organellar genomes and sex chromosomes —Insights from the case of the liverwort *Marchantia polymorpha*

By Kanji Ohyama,[*1,*2,†] Miho Takemura,[*1] Kenji Oda,[*3] Hideya Fukuzawa,[*4] Takayuki Kohchi,[*4] Sigeki Nakayama,[*5] Kimitsune Ishizaki,[*4] Masaki Fujisawa[*6] and Katsuyuki Yamato[*4]

(Communicated by Yasuyuki Yamada, m.j.a.)

Abstract: The complete nucleotide sequence of chloroplast DNA (121,025 base pairs, bp) from a liverwort, *Marchantia polymorpha*, has made clear the entire gene organization of the chloroplast genome. Quite a few genes encoding components of photosynthesis and protein synthesis machinery have been identified by comparative computer analysis. We also determined the complete nucleotide sequence of the liverwort mitochondrial DNA and deduced 96 possible genes in the sequence of 186,608 bp. The complete chloroplast genome encodes twenty introns (19 group II and 1 group I) in 18 different genes. One of the chloroplast group II introns separates a ribosomal protein gene in a *trans*-position. The mitochondrial genome contains thirty-two introns (25 group II and 7 group I) in the coding regions of 17 genes. From the evolutionary point of view, we describe the origin of organellar introns and give evidence for vertical and horizontal intron transfers and their intragenomic propagation. Furthermore, we describe the gene organization of the Y chromosome in the dioecious liverwort *M. polymorpha*, the first detailed view of a Y chromosome in a haploid organism. On the 10 megabase (Mb) Y chromosome, 64 genes are identified, 14 of which are detected only in the male genome. These 14 genes are expressed in reproductive organs but not in vegetative thalli, suggesting their participation in male reproductive functions. These findings indicate that the Y and X chromosomes share the same ancestral autosome and support the prediction that in a haploid organism essential genes on sex chromosomes are more likely to persist than in a diploid organism.

Keywords: chloroplast genome, mitochondrial genome, sex chromosomes, *Marchantia polymorpha*, *trans*-splicing, evolution of sex chromosomes

I. Gene organization of the liverwort chloroplast genome

Since the presence of chloroplast DNA was first detected in the chloroplasts of *Chlamydomonas reinhardtii*, the molecular aspects of chloroplast DNA have been intensively studied in various species of green organisms. Chloroplasts are photosynthetic organelles that have their own genetic system, separate from the nuclear genome which also encodes a number of chloroplast proteins. The organization of chloroplast genomes has been reviewed elsewhere.[1),2)] The complete nucleotide sequences of the chloroplast genomes of a liverwort, *Marchantia polymorpha* (Fig. 1),[3)–6)] of angiosperms

[*1] Laboratory of Plant Molecular Biology, Research Institute for Bioresources and Biotechnology, Ishikawa Prefectural University, Ishikawa, Japan.
[*2] Recipient of Japan Academy Prize in 2008.
[*3] Research Institute for Biological Sciences, Okayama, Japan.
[*4] Laboratory of Plant Molecular Biology, Graduate School of Biostudies, Kyoto University, Kyoto, Japan.
[*5] Plant Genetic Engineering Research Unit, Division of Plant Sciences, National Institute of Agrobiological Sciences, Ibaraki, Japan.
[*6] Central Laboratories for Frontier Technology, Kirin Holdings Co., Ltd., Ishikawa, Japan.
[†] Correspondence should be addressed: K. Ohyama, 1-308, Suematsu, Nonoichi-machi, Ishikawa 921-8836, Japan (e-mail: kohyama@ishikawa-pu.ac.jp).

doi: 10.2183/pjab/85.108

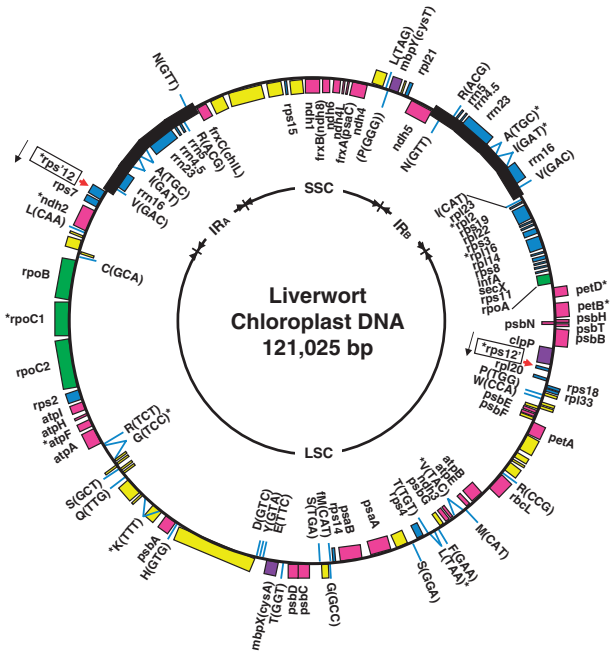

Fig. 1. Revised genetic map of the chloroplast genome of the liverwort *Marchantia polymorpha*. IRA, IRB, SSC and LSC on the inner circle indicate the inverted repeat regions, the small single-copy region and the large single-copy region, respectively. Genes shown inside the map are transcribed clockwise, and those outside are transcribed anticlockwise. Asterisks indicate genes with introns. Genes for tRNAs are indicated by the one-letter amino acid code with the unmodified anticodon. Identified protein genes and rRNA genes are indicated by gene symbols, and the remaining open boxes represent unidentified ORFs. Red arrows indicate the sites of the *rps12 trans*-splicing gene. Genes are color coded according to their functions: Photosynthesis and electron transport (*rbc*, *psa*, *psb*, *pet*, *ndh*, *atp*, *frx*); Transcription (*rpo*); Translation (*rpl*, *rps*, *rrn*, *trn*); Miscelleaneous (*mbp*, *chl*, *clp*); Unidentified ORF.[3)–6)]

such as tobacco (*Nicotiana tabacum*),[7)] rice (*Oryza sativa*)[8)] and pine (*Pinus thunbergii*),[9)] and of a protozoan *Euglena gracilis*[10)] have been reported, providing new knowledge of chloroplast genome organization and gene expression.

1. Chloroplast genome size and physical maps. The chloroplast genomes from several species of non-angiosperm land plants have been described. Chloroplast DNA of the protozoan flagellate *E. gracilis* forms circular molecules of 40 µm in contour length as shown by electron microscopy.[11)] The chloroplast DNA of *C. reinhardtii* consists of 62 µm circular molecules.[12)]

All chloroplast DNAs reported so far range in terms of their coding capacity between 120 and 160 kilobase pairs (kb). Sizes of the chloroplast DNA from the ferns *Asplenium nidus* and *Pteris vittata* resemble those of higher plants.[2)] Chloroplast DNA has been isolated and characterized from several species of bryophytes. Chloroplast DNA from a liverwort, *Sphaerocarpos donnellii*, consists of circular molecules 38.5 µm long. The liverwort, *M. polymorpha*, has chloroplasts with circular DNA molecules of 32 µm in length, and the physical map has been constructed by digestion with the restriction endonucleases *Sal*I, *Bam*HI, *Kpn*I, and *Xho*I.[13)] Genome sizes of chloroplast DNAs obtained from *E. gracilis*, *C. reinhardtii*, liverworts, and ferns do not necessarily represent the minimum size of DNA needed to function in plastids. In fact, the chloroplast genome from the liverwort *M. polymorpha* carries additional genes that are not present in the chloroplast genomes of higher plants.[6)]

2. Chloroplast ribosomal RNA(rRNA) genes (*rrn* operon). Several differences between land plants and green algae in the *rrn* operon are worth noting here. Chloroplast ribosomes generally are 70S prokaryotic ribosomes sharing similarity with those of *Escherichia coli*. A specific feature of the liverwort chloroplast ribosomes is the presence of four species of rRNAs, namely the 16S, 23S, 5S, and 4.5S rRNAs.[3),14),15)] The nucleotide sequence of the 4.5S rRNA in these chloroplasts corresponds to the 3′ terminal portion of the bacterial 23S rRNA. Therefore, the RNA components of chloroplast ribosomes are equivalent to those of *E. coli*. While the *rrn* operons between angiosperm plants and bryophytes show the same gene organization, the chloroplast *rrn* operon of *E. gracilis* wild-type strain Z has three species of rRNAs, 16S, 23S, and 5S rRNA which are similar to those of *E. coli*. However, the *Euglena* chloroplast genome has three complete sets of the *rrn* operon and one additional 16S rRNA (called the supplementary 16S rRNA; s16S rRNA) gene.[10)]

3. Chloroplast transfer RNA (tRNA) genes and codon usage. Transfer RNA genes for 31 different tRNA species have been detected in the liverwort chloroplast genome (Table 1).[4)] Of these, 5 tRNA genes are present as duplicates in the inverted repeat (IR) regions. Consequently, the liverwort chloroplast genome has 36 tRNA genes in addition to a pseudogene for proline tRNA(GGG) in the small single-copy (SSC) region. The genes for these tRNAs are scattered over the genome. Six

Table 1. Codon table and unmodified anticodons of tRNAs coded by the liverwort chloroplast genome

Codon	Anticodon		Codon	Anticodon		Codon	Anticodon		Codon	Anticodon	
UUU, UUC	Phe	GAA*1	UCU, UCC	Ser	GGA	UAU, UAC	Tyr	GUA	UGU, UGC	Cys	GCA
UUA, UUG	Leu	UAA, CAA	UCA, UCG		UGA	UAA, UAG	ter		UGA	ter	
									UGG	Trp	CCA
CUU, CUC, CUA, CUG	Leu	UAG	CCU, CCC, CCA, CCG	Pro	*GGG*, UGG	CAU, CAC	His	GUG	CGU, CGC, CGA, CGG	Arg	ACG, CCG
						CAA, CAG	Gln	UUG			
AUU, AUC	Ile	GAU	ACU, ACC, ACA, ACG	Thr	GGU, UGU	AAU, AAC	Asn	GUU	AGU, AGC	Ser	GCU, UCU
AUA		CAU				AAA, AAG	Lys	UUU	AGA, AGG	Arg	
AUG	Met/fMet	CAU*3									
GUU, GUC, GUA, GUG	Val	GAC, UAC	GCU, GCC, GCA, GCG	Ala	UGC*2	GAU, GAC	Asp	GUC	GGU, GGC, GGA, GGG	Gly	GCC, UCC
						GAA, GAG	Glu	UUC			

The AUG codon is an initiation codon. Termination codons (UAA, UAG and UGA) are indicated by ter. Amino acids are shown by three-letter symbols. The proline tRNA(GGG) gene shown in italic is a pseudogene.

*1 The phenylalanine tRNA(GAA) reads the UUU codon by wobbling (G-U).

*2 The alanine tRNA(UGC) reads the GCU, GCC and GCG codon by expanded wobbling (U-U, U-C and U-G, respectively).

*3 The initiation codon (AUG) is read by the formylmethionine tRNA(CAU), and the methionine codon by the methionine tRNA(CAU).[4)]

tRNA genes are split by an intron. No tRNA molecule needs to be imported from the cytoplasm to the chloroplasts, since the 31 species of tRNAs deduced from the DNA sequence are sufficient to decode all of the universal codons provided that some codons can be recognized by wobbling (G-U) or expanded wobbling (U-N, two out of three recognition). However, the possibility of tRNA transport from cytoplasm to chloroplasts cannot be excluded, since mitochondria in higher plants import several species of tRNA molecules from the cytoplasm as described below. The number of tRNA species in chloroplasts is much smaller than the over 50 species in *E. coli*, but higher than the 24 in the yeast mitochondrial genome and the 22 in human mitochondria, however, in the latter two mitochondrial genomes the codon table used deviates from the universal one. For tobacco and rice chloroplasts, 30 species of tRNA genes have been reported in each.[7),8)]

4. Expression of overlapping genes for photosynthesis. From the complete nucleotide sequence of the liverwort chloroplast DNA, two open reading frames, *psbN* and *psbT*, have been deduced to be located between the *psbB* and *psbH* genes. In particular, the liverwort *psbN* transcripts overlap on the opposite DNA strand with the *psbB* gene and both are actively transcribed in liverwort as well as in pea chloroplasts. Consequently the *psbN* transcripts are partially complementary to the primary transcripts of the *psbB* operon. These observations imply a possibility for controlled mRNA processing or premature transcription termination in the *psbB* operon (Fig. 2).[16)] The products of both the *psbB* and the *psbN* genes have been identified as components of the PSII complex in chloroplasts.[17)] This may be the first observation of dual functions of a chloroplast gene, one being a regulatory function by antisense RNA and the other encoding a structural component of the PSII complex. Gene clusters are also formed by the ATP synthase subunit genes *atpB-atpE* and *atpI-atpH-atpF-atpA*, respectively in the liverwort chloroplast genome.[18)]

5. The *rbcL* gene coding for the ribulose-1,5-bisphosphate carboxylase/oxygenase (large subunit, LS). The chloroplast enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase, which catalyzes the fixation of CO_2, consists of eight

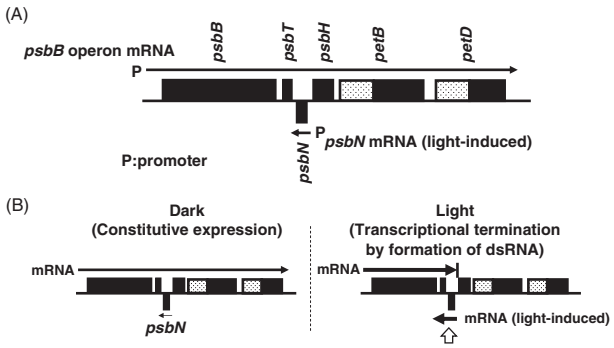

Fig. 2. Transcriptional regulation by the formation of double stranded RNA. (A) The *psbN* gene is found on the opposite DNA strand from the *psbB* gene. (B) In the dark, transcription occurs from the *psbB* gene to the *petD* gene (left). In the light, transcription of the *psbN* gene is induced and its transcripts can form a double stranded RNA with the mRNAs of the *psbB* operon, resulting in the inhibition of transcription from the *psbH* gene to the *petD* gene. Closed boxes and hatched boxes are exons and introns, respectively.[16] ⇧ A region of double strand RNA formation.

identical large and eight identical small subunits encoded in the chloroplast and nuclear genomes, respectively. The *rbcL* genes in *C. reinhardtii* and *E. gracilis* have been located on physical maps of their chloroplast DNAs. The regions surrounding the *rbcL* genes have different gene organizations in *C. reinhardtii*,[19] *E. gracilis*,[20] and liverwort,[3] the organization of the liverwort gene around the *rbcL* gene rather being similar to that in higher plants.[7),8)]

6. Genes for subunits of NADH-ubiquinone reductase. The mitochondrial NADH-ubiquinone reductase is an assembly of more than 20 different subunits. Seven of these subunits, ND1, ND2, ND3, ND4, ND4L, ND5, and ND6, are encoded in the human mitochondrial genome.[21] Interestingly, homologues of these genes (*ndh1*, *ndh2*, *ndh3*, *ndh4*, *ndh4L*, *ndh5*, and *ndh6*), and genes for the additional subunits *ndh7*, *ndh8*, and *ndh9* have been identified in the liverwort chloroplast genomes. The function of these genes is believed to be another electron transport system in chloroplasts.[22]

7. Newly found genes in the liverwort chloroplast genome. In the liverwort chloroplast genome, there are three open reading frames, designated the *frxA*(*psaC*), *frxB*(*ndh8*) and *frxC*(*chlL*) genes, that show amino acid sequence similarity with an iron-sulfur (4Fe-4S) protein found in several microorganisms.[23] Two repeated sequences Cys-X-X-Cys-X-X-Cys-X-X-X-Cys-Pro-, which are the characteristic repeat units of 4Fe-4S ferredoxin, are present in the *frxA*(*psaC*) gene product. The *frxA*(*psaC*) gene product is found in the chloroplast PSI complex, and has been suggested to be an apoprotein for the iron-sulfur center A, B, or both.[24] The *frxB*(*ndh8*) gene product also contains nine cysteine residues like the *frxA*(*psaC*) product, its amino acid sequence can be aligned with that of bacterial 4Fe-4S type ferredoxin and it is a subunit of the NDH complex. The *frxC*(*chlL*) gene product has similarity to the bacterial nitrogenase component encoded by the *nifH* gene, the Fe-protein. The *frxC*(*chlL*) gene encoded protein has nine cysteine residues, four of which are located in the region of similarity to the *nifH* gene products. Curiously, no gene corresponding to the *frxC*(*chlL*) gene has been found in the tobacco or rice chloroplast genomes. The *frxC*(*chlL*) gene product also shows high similarity with a putative protein encoded by the F202 gene of a purple non-sulfur bacterium, *Rhodopseudomonas capsulata*.[25] We have found a homologue of the *frxC*(*chlL*) gene in the cyanobacterium *Synechocystis* PCC6803[26] and its participation in the biosynthesis of chlorophyll has been shown in *Plectonema boryanum*.[27]

The two liverwort open reading frames coding for polypeptides of 370 and 288 amino acids have been designated *mbpX*(*cysA*) and *mbpY*(*cysT*), respectively. These gene products have telling similarity with those of the *hisP* and *hisQ* gene products of the histidine transport system in *Salmonella typhimurium*, and those of the corresponding respective *malK* and *malF* gene products in the inner membrane complex of the maltose transport system in *E. coli*.[28] These gene products may be associated with components derived from the nuclear genome, forming a sulfate transport complex in chloroplasts. These genes are absent from the chloroplast genomes of tobacco[7] and rice.[8]

II. Gene organization of the liverwort mitochondrial genome

Plant mitochondrial genomes (mtDNAs) are variable in size, ranging from about 200 kb in *Brassica* to over 2,000 kb in muskmelon, and are more complex than those of mammalian and fungal mitochondria.[29)–32)] Moreover, most plant mtDNAs have a complex multipartite organization in which

a hypothetical master chromosome is resolved into smaller subgenomic molecules by homologous recombination between repeated sequences.[33),34)] These features hamper the determination of the complete gene organization of the mitochondrial genomes of plant species.

1. The complexity of plant mitochondrial genomes. The analysis of the gene organization and structure of flowering plant mitochondrial genomes is made difficult by their dynamic and variable structures. This complexity is mostly due to the presence of large inverted and tandem repeated sequences in mtDNA species. These repeated nucleotide sequences cause frequent homologous recombination events which produce a large number of multipartite molecules. In contrast to these genome complexities in vascular plants we found that the mtDNA from the liverwort *M. polymorpha* consists of a single species of DNA molecules. The absence of large repeated sequences and homologous recombination mechanisms in liverwort mitochondria are probably responsible for this uniformity of the mtDNA. This observation implies that homologous recombination mechanisms in mitochondrial genomes may have been introduced into angiosperm plants after the branching of the angiosperm plants from the bryophytes. Secondary loss of recombination may be responsible for the homogeneous mitochondrial genome in the white mustard, *Brassica hirta* which contains a single species of mtDNA of 208 kb in length.[35)] This mitochondrial genome probably lost homologous recombination mechanisms late in evolution, because in the mtDNAs of several closely related *Brassica* species large repeated sequences and frequent recombination events are observed.

2. Mitochondrial genome size and gene composition. Analysis of the liverwort mtDNA by electron microscopy and restriction endonuclease mapping indicated that the liverwort mitochondrial genome is a single circular molecule of about 184.4 kb in size.[36)] In the complete sequence of the liverwort mtDNA (186,608 bp), we detected 96 genes.[37),38)] These include genes for three species of ribosomal RNA, 29 genes for 27 species of transfer RNA, 31 open reading frames for functionally known proteins. These are 16 ribosomal proteins, 3 subunits of the H^+-ATPase, 3 subunits of the cytochrome *c* oxidase, the apocytochrome *b* protein, and 8 subunits of the NADH ubiquinone oxidoreductase. Two pseudogenes are present, *ψcob* for cytochrome *b* and *ψnad7* for an NADH dehydrogenase subunit, respectively. The genes identified on the liverwort mitochondrial genome are summarized in Fig. 3.

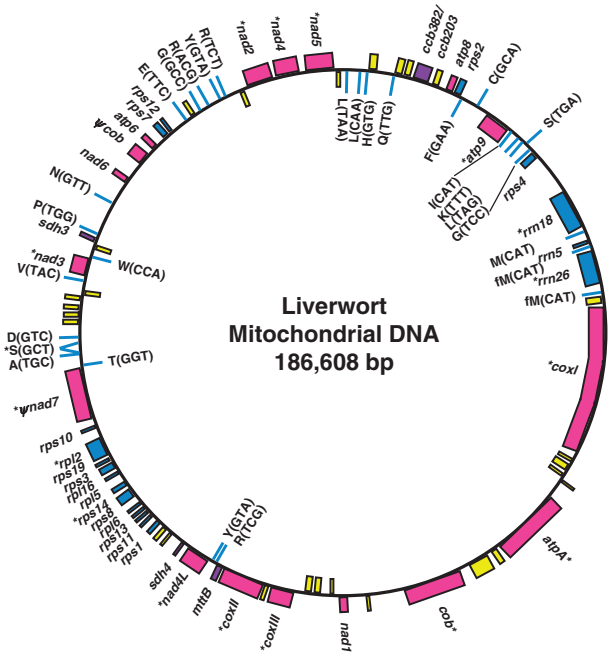

Fig. 3. Gene organization of the mitochondrial genome from the liverwort *Marchantia polymorpha*. Genes shown inside the map are transcribed clockwise, those outside are transcribed anticlockwise. Asterisks indicate genes with introns. Genes for tRNAs are indicated by the one-letter amino acid code with the unmodified anticodon. Identified protein genes and rRNA genes are indicated by gene symbols. The remaining open boxes represent unassigned ORFs. Genes are color coded according to their functions: ■ Respiration (*cox*, *cob*, *nad*, *atp*); ■ Translation (*rpl*, *rps*, *rrn*, *trn*); ■ Miscelleaneous (*sdh*, *ccb*, *mtt*); □ Unidentified ORF.[37),38)]

3. Mitochondrial transfer RNA (tRNA) genes and codon usage. We located 29 tRNA genes encoding 27 species of tRNAs, including two duplicates, *trnfM* and *trnY*.[37),39)] While the tRNA population characterized in bean mitochondria consists of mitochondrial and nuclear-encoded species, and some tRNAs from potato (*Solanum tuberosum*) are chloroplast-like species that originated from the chloroplast genome and were transferred to the mtDNA during evolution, none of the tRNA genes detected in the liverwort mitochondrial genome is chloroplast-like. However, even after allowing for wobbling and modification of the anticodons, genes for two species of tRNAs (*trnI*

Table 2. Codon table and unmodified anticodons of tRNAs coded by the liverwort mitochondrial genome

Codon	Anticodon		Codon	Anticodon		Codon	Anticodon		Codon	Anticodon	
UUU, UUC	Phe	GAA	UCU, UCC	Ser	UGA	UAU, UAC	Tyr	GUA	UGU, UGC	Cys	GCA
UUA, UUG	Leu	UAA	UCA			UAA	ter		UGA	ter	
		CAA	UCG			UAG	ter		UGG	Trp	CCA
CUU, CUC, CUA, CUG	Leu	UAG	CCU, CCC, CCA, CCG	Pro	UGG	CAU, CAC	His	GUG	CGU, CGC, CGA, CGG	Arg	ACG, UCG
						CAA, CAG	Gln	UUG			
AUU, AUC	Ile	*GAU*	ACU, ACC	Thr	AGU	AAU, AAC	Asn	GUU	AGU, AGC	Ser	GCU
AUA		CAU	ACA		*UGU*	AAA, AAG	Lys	UUU	AGA, AGG	Arg	UCU
AUG	Met/fMet	CAU	ACG								
GUU, GUC, GUA, GUG	Val	UAC	GCU, GCC, GCA, GCG	Ala	UGC	GAU, GAC	Asp	GUC	GGU, GGC, GGA, GGG	Gly	GCC, UCC
						GAA, GAG	Glu	UUC			

The AUG codon is an initiation codon. Termination codons (UAA, UAG and UGA) are indicated by ter. Amino acids are shown by three-letter symbols. The isoleucine tRNA(GAU) and threonine tRNA(UGU) indicated in italic are predicted to be imported from the nucleus. The isoleucine tRNA(GAU) and threonine tRNA(AGU) indicated by box have been shown to be imported from the nucleus.[40]

and *trnT*) can not be detected, but are required for reading all of the codons in the liverwort mitochondrial protein coding genes. To translate all codons used in liverwort mitochondrial genome, these two tRNA species must be imported from the cytoplasm into the mitochondria (Table 2).[40]

III. Characterization of organellar introns in plants

On the basis of structural features, such as nucleotide sequences or potential secondary structures, two types of introns are found in organelles, group I and group II introns[41] although an another type of introns (group III) is reported to be present in *E. gracillis* chloroplast genome.[10] Group I introns are widely distributed over the genomes of bacteriophage, prokaryotes and organelles. Group II introns are present in mitochondrial genomes of fungi and plants, and in chloroplast genomes. Group I and group II introns were originally described as two families of introns which, in addition to encoding proteins, possess unique secondary structures. Since then, the structured RNA components of group I introns have been shown to be responsible for the self-splicing phenomenon discovered by Cech,[42] while various functions are ascribed to the proteins encoded by some of these introns.

1. Chloroplast introns and their splicing. Introns are present not only in the genes of eukaryotic cells, but also in yeast mitochondrial genes and in the chloroplast 23S rRNA gene of *C. reinhardtii*.[43] Chloroplast DNA from *E. gracilis* contains a minimum of 149 introns. There are 72 individual group II introns, 46 individual group III introns, 10 group II introns and 18 group III introns that are components of twintrons.[10] In other plants, chloroplast introns belong to either group I or group II depending on their secondary structure as described for mitochondrial introns.[41] There are 20 different introns in the liverwort chloroplast genome. Only one group I intron has been found in the leucine tRNA(UAA) gene. The rest of the introns in the liverwort chloroplast genome has the typical secondary structure of group II introns.[4]

(1) *A trans-splicing gene in the liverwort chloroplast genome.* Among these split genes, as described earlier, a coding sequence corresponding to the ribosomal protein S12 gene (*rps12*) of *E. coli* is split into three exons. Exon 1 is located far from the two other exons on the opposite DNA strand (*trans-*

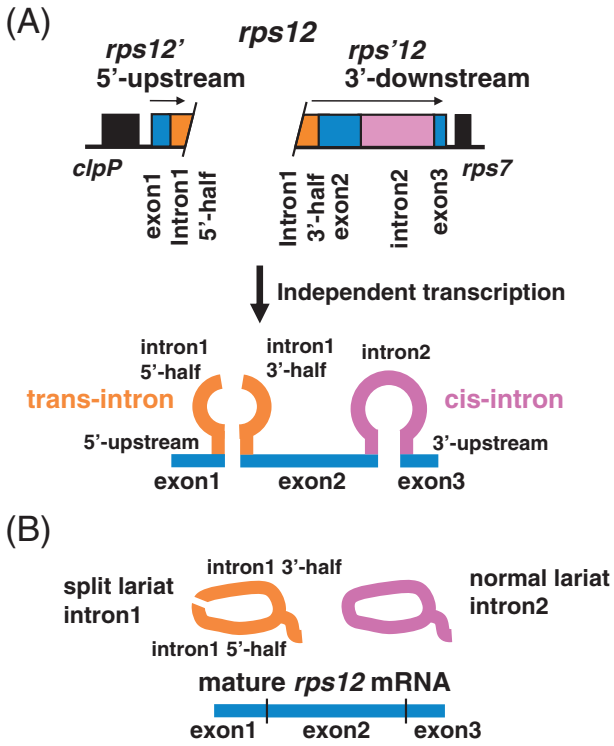

Fig. 4. *Trans*-splicing of the gene for ribosomal protein S12. (A) The 5′-upstream region and 3′-downstream region of the *rps12* gene are coded on opposite DNA strands and transcribed independently. (B) Intron 1 (*trans*) is spliced as a split lariat intron. Intron 2 (*cis*) is spliced as a normal lariat intron and a single mature mRNA is formed.[44]

split gene) (Fig. 4).[44] S1-nuclease and northern hybridization analysis with RNA prepared from liverwort chloroplasts has shown that transcripts of exon 1 and exons 2–3 of the *rps12* gene with neighboring genes are synthesized separately. There is no particular splicing order of intron 1 (*trans*) and intron 2 (*cis*). The gene (*psaA*) for the P700 chlorophyll α-apoprotein of PSI in the *C. reinhardtii* chloroplast genome is also *trans*-split into three exons with two group II introns.[45] The exons are like the *rps12* exons in the liverwort transcribed independently from different DNA strands and are *trans*-spliced into mature mRNA. Mutant analysis of *C. reinhardtii* has shown that nuclear mutants deficient in PSI are specifically affected in *psaA* RNA splicing. This study showed that gene products from the nuclear genome are involved in the chloroplast RNA splicing event that is required for expression of the *psaA* gene. Recently, RNA *trans*-splicing was also reported to play an important role in the expression of the negative RNAi regulator ERI-6/7 gene in *Caenorhabditis elegans*.[46] RNA splicing of two *cis*-introns in the liverwort *orf203*(*clpP*) gene most likely processes successively in the 5′ to 3′ direction, indicating that ordered splicing occurs in the chloroplasts of land plants.[47] However, RNA splicing of *cis*- and *trans*-introns present in the *rps12* gene in liverwort chloroplasts takes place independently. This indicates that RNA processing of transcripts is closely linked to how the order of RNA splicing is decided. To elucidate the RNA splicing mechanisms in chloroplasts, the development of an *in vitro* RNA splicing system has been attempted.[48]

2. Introns in the liverwort mitochondrial genome. The 186,608 bp of the liverwort mitochondrial genome are predicted to code for 96 genes, of which 17 are interrupted by a total of 32 introns.[37),38] Based on their sequence and structure analysis, twenty-five of the introns are assigned to the group II, the remaining seven belong to the group I. The seven group I introns are the 3rd, 4th, and 6th to 9th introns of the *cox1* gene coding for cytochrome *c* oxidase subunit 1, and the sole intron in the *nad5* gene for subunit 5 of the respiratory-chain NADH dehydrogenase. The splice junctions of the nine *cox1* introns were determined by a comparison with fungal *coxI*(*cox1*) genes by taking advantage of the presence of the consensus nucleotides at intron-exon junctions.[49]

(1) *Introns in the cox1 gene of the liverwort mitochondrial genome.* Interestingly, while the *cox1* genes of higher plants contain no introns at all, more than half of the liverwort *cox1* introns happen to be inserted at the same sites where introns have been reported to exist in the genes of fungal mitochondrial genomes (Fig. 5).[50] The site of insertion of the liverwort 2nd intron coincides with that of first intron from the yeast *Saccharomyces cerevisiae*. The liverwort 4th intron is inserted at the same site as the first intron in the fission yeast *Schizosaccharomyces pombe*. The site of the insertion of the liverwort 6th intron is identical to that of introns in three filamentous fungi, *Neurospora crassa* and *Podospora anserina*. The liverwort 7th and 8th introns interrupt the *cox1* gene at the same places as introns in yeast, fission yeast and *Podospora*. Of all these introns, only the liverwort 2nd intron and the yeast first intron belong to the group II of introns.

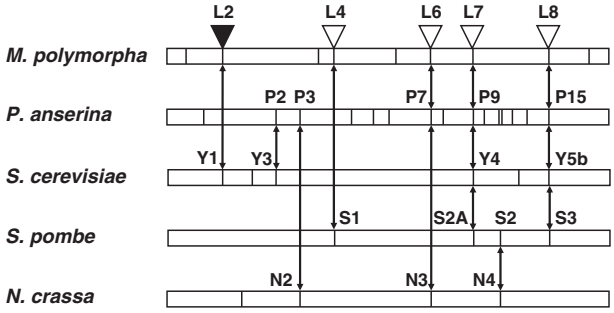

Fig. 5. Schematic alignments of *cox1* genes with the same insertion sites of introns from *M. polymorpha*(L), *P. anserina*(P), *S. cerevisiae*(Y), *S. pombe*(S), and *N. crassa*(N). Arrows indicate introns inserted at identical sites in the mitochondrial genomes of the different species. Open and filled triangles indicate group I and group II introns, respectively.[50)]

(2) *Secondary structure analysis of the liverwort group I introns.* Secondary structure models of liverwort group I introns are all typical of subgroup IB, the same group that all of the fungal introns inserted at sites identical to the liverwort 4th, 6th, 7th and 8th introns in the *cox1* gene belong to. The secondary structures and primary sequences of the 6th to 8th introns in the *cox1* gene have been compared with those of their fungal counterparts as well as a broader selection of subgroup IB introns,[49)] but these comparisons remained inconclusive.

(3) *Analysis of open reading frames in liverwort group I introns.* The two liverwort 4th and 8th introns of *cox1* gene have long open reading frames which continue in frame from the upstream exons. The two proteins translated from these reading frames exhibit characteristic dodecapeptide motifs. We searched our database of intron-encoded proteins for possible close relatives of the putative proteins encoded in their introns. In both cases, the only sequences with which any one of these *orf*s could be aligned over nearly the entire length turned out to be the fungal introns inserted at the respective homologous site, that is, the first intron in the *S. pombe cox1* gene for the liverwort 4th intron, and the 15th intron in the *P. anserina cox1* gene for the liverwort 8th intron in the *cox1* gene.

(4) *Intragenomic propagation of liverwort mitochondrial group II introns.* Six groups of highly similar group II introns are found in the liverwort mitochondrial genome.[50),51)] The well-conserved secondary structures, compensatory nucleotide changes and high percentages of nucleotide sequence identities among the group II introns in each similarity group strongly suggest that respective ancestral introns propagated intra-genomically long after the divergence of the bryophytes from the fungal kingdom. This evolutionary spreading to additional locations implies that the compatibilities between the exon binding sequence (EBS) of an intron and the intron binding sequence (IBS)-like sequence of the exon at the new location must be compatible to allow the insertion of the intron into the new mRNA sequence *via* a reverse splicing process.[50)] A mechanism of amplifying and spreading of introns *via* an RNA intermediate has been proposed.[51)] This intron acquiring processes proceeds by reverse-splicing, reverse transcription, and finally homologous recombination to integrate the intron as DNA into the genome. The compatibilities of EBS and IBS sequences in similar introns in the liverwort mitochondrial genome suggest that these intron families have arisen *via* such reverse splicing processes.[51)] The mitochondrial element which mediates the amplification and integration of RNA sequences into new sites in the genome seems to be the RNA maturase encoded in these group II introns, because amino acid sequence motifs characteristic of reverse transcriptases are found in these reading frames.[52)] There is also a likely trace of an RNA maturase encoded in the 2nd intron of *ψnad7* gene of the liverwort mitochondrial genome that had been inserted by a transposon-like mechanism. These observations suggest the mobility of the RNA maturase genes and the possibility of a participation of RNA maturases in the intron propagation.[51)]

IV. RNA editing in plant organellar genomes

RNA editing was first identified as the insertion and the deletion of uridine residues in mitochondrial mRNAs of kinetoplastids in protozoa.[53)] RNA editing by the conversion of C residues to U residues in the mRNA was subsequently reported in mitochondria and chloroplasts in several species of plants.[54),55)] RNA editing is apparently lacking in liverwort mitochondria and chloroplasts since the nucleotide sequences of these liverwort organellar DNAs are well-conserved at the DNA level for maintaining the expected amino acid sequences of evolutionarily conserved proteins (Fig. 6).[56)] Moreover, we found a very high frequency of conversions

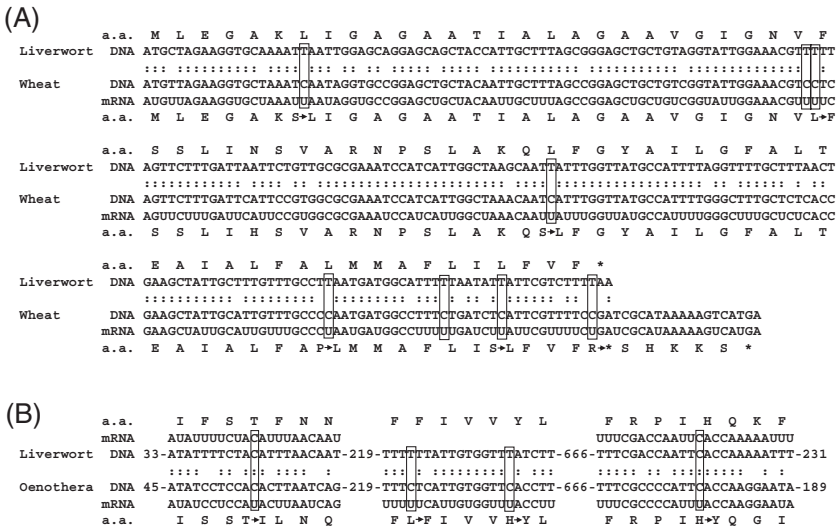

Fig. 6. No RNA editing in the liverwort mitochondrial genome. (A) RNA editing is not seen in the regions of the liverwort mitochondrial mRNA of the *atp9* gene where RNA editing was demonstrated in the wheat *atp9* gene. Boxed residues are RNA editing sites (C to U conversions) in the wheat *atp9* gene. The corresponding sites in the liverwort *atp9* gene are not required to be edited because T residues are encoded in the genomic sequence. (B) RNA/cDNA sequence analyses of the liverwort *cob* gene demonstrated the lack of RNA editing at any potential RNA editing site in the liverwort *cob* gene where RNA editing has been reported in higher plants. Boxed residues demonstrate RNA editing sites in *Oenothera cob* gene.[56]

from T to C residues in the mitochondrial genes of higher plants in comparison to the corresponding liverwort genes. This implies that the RNA editing system was introduced into the branch of the angiosperm plants after the divergence from the bryophytes as a mechanism for reverting the frequent conversion of T residues to C residues, rather than for the conservation of amino acid sequences.[57]

1. Evolutionary mechanisms for the restoration of AT→GC changes in angiosperms. It has previously been proposed that the CGG codon is specifying tryptophan instead of arginine in plant mitochondria.[58] Gualberto *et al.*[54] and Covello and Gray[59] described RNA editing with C→U conversions (C residues that are edited to U residues in the corresponding mRNA sequences) in the *coxII* (cytochrome *c* oxidase, subunit II) mRNA in wheat mitochondria. These conversions result in UGG codons in the mRNA that can be recognized by the tryptophan tRNA according to the universal genetic code, and it was accordingly suggested that RNA editing contributes to the conservation of mitochondrial protein sequences in plants. We have looked for the requirement of RNA editing in the sequence of the liverwort mtDNA at positions in genes where edited RNA was reported in higher plants, but almost all C residues in RNA editing sites in higher plants are coded as T residues in the liverwort mtDNA (Fig. 6A). Two possible sites were found in the liverwort *cob* gene, but RNA/cDNA sequencing did not show any evidence of RNA editing (Fig. 6B).[50] These results indicate that liverwort mitochondria do not require RNA editing for gene conversion.

From an evolutionary perspective, there are at least two possibilities when RNA editing appeared to correct protein coding sequences: (1) RNA editing was an ancient mechanism in the RNA world or (2) RNA editing evolved after the evolutionary divergence of the angiosperm plants. The evidence described here strongly suggests that RNA editing is not a relic of an ancient mechanism for protein sequence conservation. It is likely that RNA editing first appeared in land plants after the divergence of angiosperm plants from non-angiosperm plants (e.g. the bryophytes). However, this does not rule out the possibility that liverwort mitochondria did use RNA editing prior to the divergence of angiosperms and bryophytes and that Marchantiid subsequently ceased to use this process.[60]

Appearance of RNA editing in angiosperm mitochondria is possibly correlated with with an

AT→GC exchange pressure in evolution. The GC content of the *coxII* gene in angiosperm plants is with e.g. 40% in wheat higher than in non-plant species (26% in yeast) and in the liverwort with 35%. This indicates that during evolution of the land plants the mutations T→C and A→G took place at higher rates in mitochondrial genomes of angiosperm plants than in the liverwort and in non-plant species.[57] In summary these results strongly suggest that RNA editing (C→U and G→A conversion) has been introduced into the mitochondrial genome of angiosperm plants to accommodate the T→C changes introduced by the AT→GC change pressure by restoring the codons required inconserved protein sequences.

V. Evolution of organellar genomes

An autogenous origin of chloroplasts and mitochondria had long been believed to have occurred within eukaryotic cells by the formation of membrane compartments. However, molecular data accumulated by DNA sequence analysis of gene organization and composition strongly suggest that chloroplasts and mitochondria originated from common ancestors with cyanobacteria and alpha-proteobacteria, respectively. We have deduced a single origin of chloroplasts based on the complete DNA sequence analysis of chloroplast genomes from several species of plant cells.[3),7)–9)] On the other hand, there is still doubt whether a single endosymbiotic event or several were involved in the establishment of the mitochondria.

1. Single origin of the chloroplast genome. From the evolutionary point of view, we have proposed a single origin of chloroplasts in green land plants.[61),62)] Analysis of chloroplast gene organization, gene sequence, gene expression system supports the idea that chloroplasts originate in evolution from specific ancestors common with cyanobacteria by a process of endosymbiosis into a progenitor cell of photosynthetic eukaryotes.[63)] During the establishment of an ancestral plant chloroplast, many prokaryotic genes must have migrated from the genome of the endosymbiotic chloroplast progenitor to the host nucleus, and many of them remain detectable as such in the nuclear genomes in present day plants. The allocation of the genes for all the chloroplast components to either the chloroplast or the nuclear genome is basically identical in both liverwort and higher plants. Therefore the present nuclear-chloroplast relationship was most likely established about 300–400 million years ago. From this single origin, the variations in the present day chloroplast genomes in plants and rearrangements in chloroplast DNA have evolved as a result of the accumulation of mutations in nucleotide sequences, but rarely from changes in gene content.

2. Evolution of the mitochondrial genome. The mitochondrial genomes of higher plants exhibit extraordinary differences in their sizes,[30)] in gene transfers from the chloroplast genome into the mitochondrial genome,[64)] and in rearrangements due to homologous recombination.[33),34)] These observations make it difficult to speculate on the ancestral mitochondrial genome.

In higher plant mitochondria only one secondary invasion event of group I introns has yet been reported.[65)] Assuming group I introns are indeed originally absent from the mitochondrial compartment of these organisms, two contrasting explanations may account for their presence in the mitochondrial genes of the liverwort, *M. polymorpha*. Either group I introns were present in the last common ancestor of bryophytes and angiosperms, and got lost in the branch leading to the latter or they were missing from that common ancestor, and have subsequently been acquired somewhere along the branch leading to the bryophytes.

VI. Gene organization of the liverwort Y chromosomes

In many sexually dimorphic organisms, sex chromosomes play key roles in sex determination and sexual development. In mammals, for example, a female has two X chromosomes, while a male has one X and one Y chromosome (XX/XY system). The mammalian Y chromosome carries a gene that induces testis development (*e.g.* mouse *SRY*)[66),67)] and thus determines male sex. In humans, the Y chromosome is smaller than most of the other chromosomes and harbors less than 200 genes, whereas the X chromosome contains 1,098 genes.[68)] The Y chromosome has accumulated mutations and lost many genes, due to suppressed recombination with the X chromosome as reviewed and discussed by Charlesworth and Charlesworth.[69)]

Sex determination systems in dioecious plants are thought to have evolved many times from hermaphroditic ancestors.[70)] Analogous to the

mammalian XX/XY system, in the dioecious angiosperm white campion (*Silene latifolia*) males are XY and females XX. The Y chromosome induces male development.[71] In papaya (*Carica papaya*), a dominant male-determining locus resides in a chromosomal region on one chromosome where recombination is suppressed and the DNA sequence has extensively diverged from its homologous region on the X chromosome.[72] Therefore, this papaya chromosome is considered an immature Y chromosome in an emerging XX/XY system. In dioecious sorrel, *Rumex acetosa*, a male individual has one X and two Y chromosomes, while a female has two X chromosomes. Here the presence of Y chromosomes has no influence on triggering male development, but as is the case in *Drosophila*, the X/autosomal balance determines the sex.[73),74]

Y chromosomes in haploid dioecious organisms also appear different from other chromosomes, but have been poorly studied. Under haploidy, both X and Y chromosomes should evolve in the same way.[75] The liverwort *M. polymorpha*, an extant species of the earliest land plants,[76] is dioecious, and the dominant forms in its life cycle are female and male haploid thalli. These are phenotypically identical until the female or male sexual organs differentiate. After fertilization by flagellated sperm, a diploid zygote develops into a sporophyte by mitosis, followed by meiosis to produce haploid spores which germinate and develop into the next generation thalli. The haploid thalli of both sexes can also propagate asexually by so-called gemmae.

Several bryophyte species have been reported to possess sex chromosomes,[77] in *M. polymorpha* the haploid set of chromosomes consists of eight autosomes and a single sex chromosome, an X chromosome in females ($n = 8 + X$) and a Y chromosome in males ($n = 8 + Y$). Therefore, unlike in the XX/XY system, the X and Y chromosomes in *M. polymorpha* are separated from each other during most of the life cycle. In *M. polymorpha*, the X and Y chromosomes align during meiotic metaphase but keep some distance from each other and are in the diakinesis stage separated earlier than the autosomes,[78),79] suggesting that no recombination between the X and Y chromosomes is possible. The extensive sequence analysis of the *M. polymorpha* Y chromosome reported here not only provides an in-depth view into the gene content and structure of a plant sex chromosome but also yields insights into the evolution of recombination-suppressed sex chromosomes in a haploid genome.

1. Sequencing of the *Marchantia* Y chromosome. We had previously found that the *M. polymorpha* Y chromosome (10 Mb) contains a family of unique repeats that are represented by a 2.4-kb BamHI repeat confined to a 4 Mb segment of the Y chromosome[80] designated YR1 (Y chromosome Region 1) (Fig. 7A). We now obtained the sequence of YR1 by sequencing 28 PAC clones that collectively cover YR1, since the accumulation of the 2.4-kb BamHI repeat family made the construction of contigs impractical. The total sequence obtained amounts to 3,200,899 bp.

For sequence analysis of the other 6 Mb segment of the Y chromosome YR2 (Y chromosome Region 2), 59 tiled PAC clones were selected from two contigs of aligned PAC clones, Contig-A and Contig-B, that cover YR2. Three PCR amplified DNA fragments filled mapping and tiling gaps in Contig-B. The total lengths of the sequences obtained are 3,467,261 bp for Contig-A and 2,530,874 bp for Contig-B, which account for more than 95% of their size as estimated from the physical mapping (Fig. 7B).[81]

Contig-A and Contig-B were cytologically mapped on the Y chromosome by fluorescence *in situ* hybridization (FISH).[82] One of the signals for the end of Contig-A terminated by clone pMM23-431A8 was detected in the immediate vicinity of the more condensed YR1. The signal for the end of Contig-B terminated by clone pMM23-359F1 was detected in the central region of YR2. This result aligns YR1, Contig-A and Contig-B on the Y chromosome in the order of YR1 — Contig-A — Contig-B.[82]

2. Genes on the *Marchantia* Y chromosome. Similarity searches against the public sequence databases and *M. polymorpha* ESTs detected 64 genes, 9 in YR1 and 55 in YR2 (Fig. 7C).[82] Genomic PCR and/or Southern blotting shows for 25 of the Y-chromosome genes similar sequences also in female DNA, while the remaining 39 genes appear to be present only in the male DNA. Fourteen genes, one in YR1 and 13 in YR2, are unique to the male genome and in addition show sexual organ-specific expression. These genes are thus candidates for male reproductive functions.

Although no closely similar homologs of the *M. polymorpha* Y-linked genes were found in

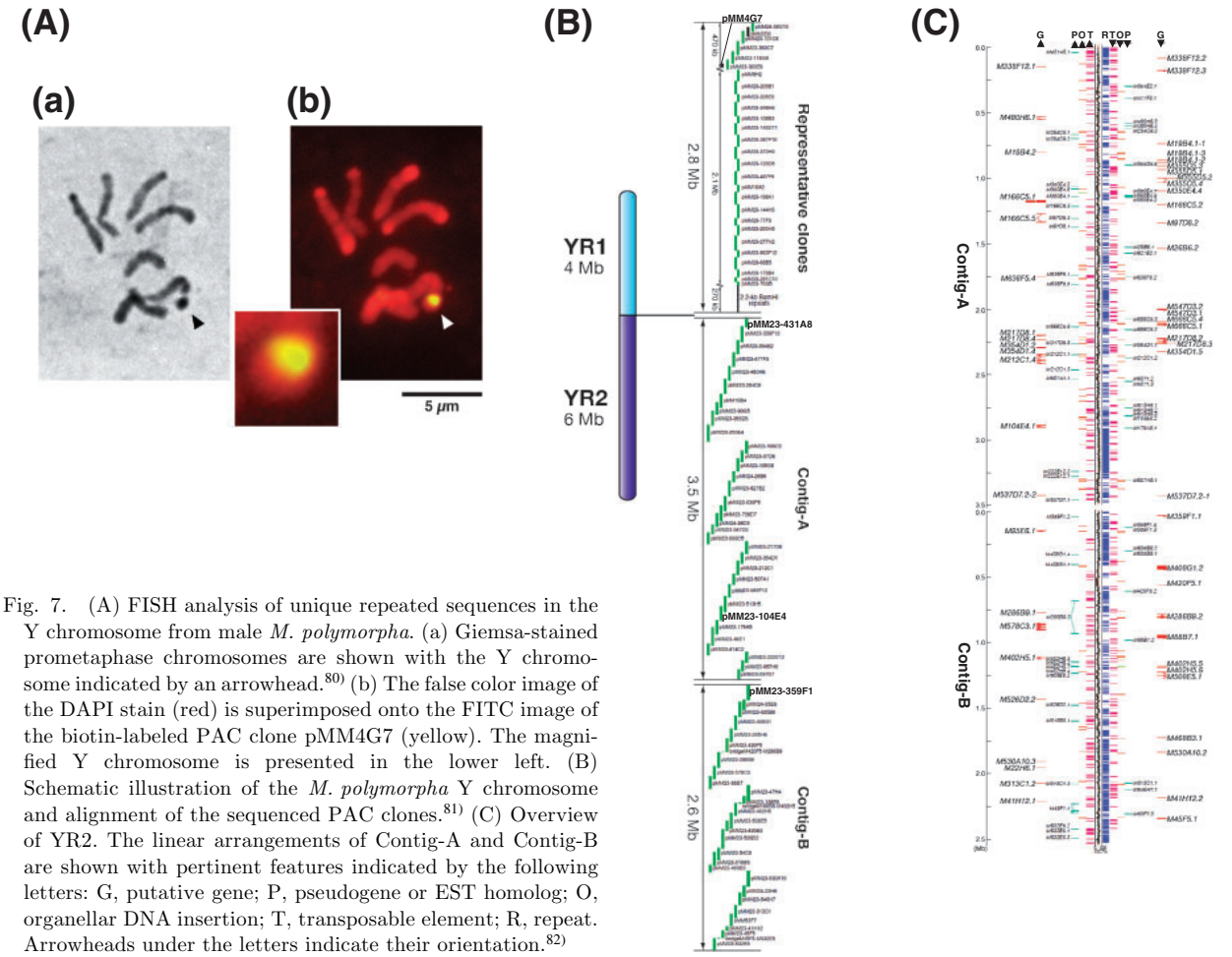

Fig. 7. (A) FISH analysis of unique repeated sequences in the Y chromosome from male *M. polymorpha*. (a) Giemsa-stained prometaphase chromosomes are shown with the Y chromosome indicated by an arrowhead.[80] (b) The false color image of the DAPI stain (red) is superimposed onto the FITC image of the biotin-labeled PAC clone pMM4G7 (yellow). The magnified Y chromosome is presented in the lower left. (B) Schematic illustration of the *M. polymorpha* Y chromosome and alignment of the sequenced PAC clones.[81] (C) Overview of YR2. The linear arrangements of Contig-A and Contig-B are shown with pertinent features indicated by the following letters: G, putative gene; P, pseudogene or EST homolog; O, organellar DNA insertion; T, transposable element; R, repeat. Arrowheads under the letters indicate their orientation.[82]

searches of other species' Y chromosomes, we identified some putative genes from open reading frames whose translated sequences resemble some animal male-fertility proteins. Among the 14 putative male reproductive genes, six encode proteins whose homologs are found in animals but not in angiosperms. Since in bryophytes male gametes are flagellated sperm, spermatogenesis in *M. polymorpha* and animals could very well share some proteins of common evolutionary descent, and these six genes may be involved in analogous functions in spermatogenesis.

Another 40 genes on the Y chromosome are expressed in vegetative thalli as well as in male sexual organs and thus may code for functions not related to male sexual differentiation. The rest of ten genes were not expressed in thalli as well as in sexual organ.

3. Repeats and transposable elements. Both YR1 and YR2 are rich in repeats, but the origins of these repeats are strikingly different. The YR1 domain consists of unique small repeat sequences of only several hundred nucleotides, which are assembled in various stoichiometries to form different arrangements of the 2.4-kb BamHI repeat family. Structures and sequences of these elements have been analyzed previously.[83] When DNA of clone pMM23-104E4 (Contig-A of YR2), which is one of the first clones isolated from YR2, and of a fragment of the 2.4-kb BamHI repeat were simultaneously applied to prometaphase chromosomes of male plants in FISH experiments, the signal of the 2.4-kb BamHI repeat identifies only YR1, the signal of pMM23-104E4 lights up only YR2 in its entirety.[84]

On the other hand, the *M. polymorpha* X chromosome (20 Mb) carries a large cluster of

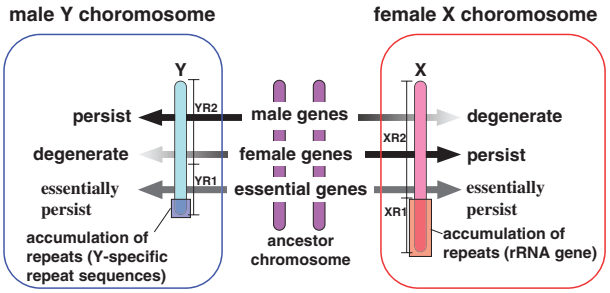

Fig. 8. Molecular evolution of the liverwort sex chromosomes. During evolution of the X and Y chromosomes the genes essential for the maintenance of each sex were sustained or degenerated as required in each sex. In parallel, repeats specific for each sex chromosome accumulated.[75),82)] Both X and Y chromosomes retained the genes coding for essential metabolic or regulatory functions required for normal growth and development.

ribosomal DNA (45S rDNA) designated XR1 (X chromosome Region 1) (Fig. 8), while no rDNA is found on the Y chromosome.[85)] When DNA of pMM23-104E4 and a fragment of the rDNA sequence on the X chromosome[85)] were used in a FISH analysis of female chromosomes, the signal of pMM23-104E4 was detected on a segment of the X chromosome (designated X chromosome Region 2; XR2) (Fig. 8) distinct from the rDNA signals, as well as on autosomes. The Contig-A region in YR2 of the Y chromosome thus contains sequences similar to motifs on a specific segment of the X chromosome as well as on autosomes.

Several repeats in YR2 are related to transposable elements. While the total complexity of common interchromosomal repeats remains unknown without sequence information of the entire genome, intrasegmental repeats of at least 200 bp and at least 90% identity alone account for 43% of YR2. An interesting similarity to YR2 is seen in humans, where common repeats such as LINE1 (long interspersed nuclear element 1), Alu and retroviral elements account for 47% of the euchromatic MSY (Male-Specific region of the Y chromosome).[86)]

4. Evolution of the *M. polymorpha* sex chromosomes. Like X and Y chromosomes of other organisms, the *M. polymorpha* sex chromosomes probably originated from a regular autosome. A widely accepted scheme for the evolution of Y chromosomes in the XX/XY system consists of three major events: acquisition of the sex-determining loci, suppression of recombination, and genetic degeneration driven by evolutionary processes such as Muller's ratchet.[69),87)] However, Bull predicts that the evolution of sex chromosomes in a haploid system is different from that in a diploid system.[75)] In a haploid organism, degeneration should not occur, because it would impair essential genes. Essential genes on the Y chromosome should thus also be present on the X chromosome, since females will likewise require these genes. Although the high synonymous substitution rate observed among the X-Y gene pairs suggests that the *M. polymorpha* Y chromosome has been long established, the 14 putative male reproductive genes are far fewer than the 40 putative general function genes expressed in thalli and sexual organs. The higher proportion of putative general function genes on the *M. polymorpha* Y chromosome is consistent with the prediction that in a predominant haploid life style degeneration must not impair non-redundant genes whose expression is essential to survival of the organism (Fig. 8).

Acknowledgments

We particularly thank Professor Emeritus Yasuyuki Yamada for his warm and generous encouragement during the time of our research work at Kyoto University. We also thank all the collaborators who contributed to the studies on liverwort genome projects. Finally we thank Professor Axel Brennicke for reading the manuscript.

References

1) Whitfeld, P.R. and Bottomley, W. (1983) Organization and structure of chloroplast genes. Annu. Rev. Plant Physiol. **34**, 279–310.
2) Palmer, J.D. (1985) Comparative organization of chloroplast genomes. Annu. Rev. Genet. **19**, 325–354.
3) Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S. *et al.* (1986) Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature **322**, 572–574.
4) Ohyama, K., Fukuzawa, H., Kohchi, T., Sano, T., Sano, S., Shirai, H. *et al.* (1988) Structure and organization of *Marchantia polymorpha* chloroplast genome. I. Cloning and gene identification. J. Mol. Biol. **203**, 281–298.
5) Ohyama, K., Kohchi, T., Fukuzawa, H., Sano, T., Umesono, K. and Ozeki, H. (1988) Gene organization and newly identified groups of genes of the chloroplast genome from a liverwort, *Marchantia polymorpha*. Photosynth. Res. **16**, 7–22.

6) Ohyama, K., Kohchi, T., Sano, T. and Yamada, Y. (1988) Newly identified groups of genes in chloroplasts. Trends in Biochem. Sci. **13**, 19–22.
7) Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T. *et al.* (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J. **5**, 2043–2049.
8) Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M. *et al.* (1989) The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evoltion of the cereals. Mol. Gen. Genet. **217**, 185–194.
9) Wakasugi, T., Tsudzuki, J., Ito, S., Nakashima, K., Tsuzuki, T. and Sugiura, M. (1994) Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. Proc. Natl. Acad. Sci. USA **91**, 9794–9798.
10) Hallick, R. B., Hong, L., Drager, R. G., Favreau, M. R., Monfort, A., Orsat, B. *et al.* (1993) Complete sequence of *Euglena gracilis* chloroplast DNA. Nucl. Acids Res. **21**, 3537–3544.
11) Gray, P. W. and Hallick, R. B. (1977) Restriction endonuclease map of *Euglena gracilis* chloroplast DNA. Biochemistry **16**, 1665–1671.
12) Rochaix, J. D. (1978) Restriction endonuclease map of the chloroplast DNA of *Chlamydomonas reinhardtii*. J. Mol. Biol. **126**, 597–617.
13) Ohyama, K., Yamano, Y., Fukuzawa, H., Komano, T., Yamagishi, H., Fujimoto, S. *et al.* (1983) Physical mappings of chloroplast DNA of liverwort *Marchantia polymorpha* L. cell suspension cultures. Mol. Gen. Genet. **189**, 1–9.
14) Yamano, Y., Ohyama, K. and Komano, T. (1984) Nucleotide sequences of chloroplast 5S ribosomal RNA from cell suspension cultures of the liverworts *Marchantia polymorpha* and *Jungermannia subulata*. Nucl. Acids Res. **12**, 4621–4624.
15) Yamano, Y., Kohchi, T., Fukuzawa, H., Ohyama, K. and Komano, T. (1985) Nucleotide sequences of chloroplast 4.5S ribosomal RNA from a leafy liverwort, *Jungermannia subulata*, and a thalloid liverwort, *Marchantia polymorpha*. FEBS Lett. **185**, 203–207.
16) Kohchi, T., Yoshida, T., Komano, T. and Ohyama, K. (1988) Divergent mRNA transcription in the chloroplast *psbB* operon. EMBO J. **7**, 885–891.
17) Ikeuchi, M., Koike, H. and Inoue, Y. (1989) N-terminal sequencing of low-molecular-mass components in cyanobacterial photosystem II core complex. Two components correspond to unidentified open reading frames of plant chloroplast DNA. FEBS Lett. **253**, 178–182.
18) Cozens, A. L., Walker, J. E., Phillips, A. L., Huttly, A. K. and Gray, J. C. (1986) A sixth subunit of ATP synthase, an F_0 component, is encoded in the pea chloroplast genome. EMBO J. **5**, 217–222.
19) Dron, M., Rahire, M. and Rochaix, J. D. (1982) Sequence of the chloroplast DNA region of *Chlamydomonas reinhardtii* containing the gene of the large subunit of ribulose bisphosphate carboxylase and parts of its flanking genes. J. Mol. Biol. **162**, 775–793.
20) Hallick, R. B. and Buetow, D. E. (1989) Chloroplast DNA. *In* The Biology of *Euglena*, Vol. IV (ed. Butetow, D. E.). Academic Press, London, pp. 351–414.
21) Chomyn, A., Cleeter, M. W., Ragan, C. I., Riley, M., Doolittle, R. F. and Attardi, G. (1986) URF6, last unidentified reading frame of human mtDNA, codes for an NADH dehydrogenase subunit. Science **234**, 614–618.
22) Shikanai, T., Endo, T., Hashimoto, T., Yamada, Y., Asada, K. and Yokota, A. (1998) Directed disruption of the tobacco *ndhB* gene impairs cyclic electron flow around photosytem I. Proc. Natl. Acad. Sci. USA **95**, 9705–9709.
23) Minami, Y., Wakabayashi, S., Wada, K., Matsubara, H., Kerscher, L. and Oesterhelt, D. (1985) Amino acid sequence of a ferredoxin from themoacidophilic archaebacterium, *Sulfolobus acidocaldarius*. Presence of an N^6-monomethyllysine and phyletic consideration of archaebacteria. J. Biochem. **97**, 745–753.
24) Oh-oka, H., Takahashi, Y. and Matsubara, H. (1989) Topological considerations of the 9-kDa polypeptide which contains centers A and B, associated with the 1- and 19-kDa polypeptides in the photosystem I complex of spinach. Plant Cell Physiol. **30**, 869–875.
25) Hearst, J. E., Alberti, M. and Doolittle, R. F. (1985) A putative nitrogenase reductase gene found in the nucleotide sequences from the photosynthetic gene cluster of *R. capsulata*. Cell **40**, 219–220.
26) Ogura, K., Takemura, M., Oda, K., Yamatao, K., Ohta, E., Fukuzawa, H. *et al.* (1992) Cloning and nucleotide sequence of a frxC-ORF469 gene cluster of Synechocystis PCC6803: Conservation with liverwort chloroplast frxC-ORF465 and nif operon. Biosci. Biotech. Biochem. **56**, 788–793.
27) Fujita, Y., Matsumoto, H., Takahashi, Y. and Matsubara, H. (1993) Identification of a nifDK-like gene (ORF467) involved in the biosynthesis of chlorophyll in the cyanobacterium *Plectonema boryanum*. Plant Cell Physiol. **34**, 305–314.
28) Higgin, C. F., Hiles, I. D., Salmond, G. P. C., Gill, D. R., Downie, J. A., Evans, I. J. *et al.* (1986) A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria. Nature **323**, 448–450.
29) Leaver, C. J. and Gray, M. W. (1982) Mitochondrial genome organization and expression in higher plants. Annu. Rev. Plant Physiol. **33**, 373–402.
30) Lonsdale, D. M. (1984) A review of the structure and organization of the mitochondrial genome of higher plants. Plant Mol. Biol. **3**, 201–206.
31) Newton, K. J. (1988) Plant mitochondrial genomes: organization, expression and variation. Annu.

Rev. Plant Physiol. Plant Mol. Biol. **39**, 503–532.
32) Levings III, C. S. and Brown, G. G. (1989) Molecular biology of plant mitochondria. Cell **56**, 171–179.
33) Lonsdale, D. M., Hodge, T. P. and Fauron, C. M.-R. (1984) The physical map and organization of the mitochondrial genome from the fertile cytoplasm of maize. Nucl. Acids Res. **12**, 9249–9261.
34) Palmer, J. D. and Shields, C. R. (1984) Tripartite structure of the *Brassica campestris* mitochindrial genome. Nature **307**, 437–440.
35) Folkerts, O. and Hanson, M. R. (1991) The male sterility-associated *pcf* gene and normal *atp*9 gene in *Petunia* are located on different mitochondrial DNA molecules. Genetics **129**, 885–895.
36) Oda, K., Kohchi, T. and Ohyama, K. (1992) Mitochondrial DNA of *Marchantia polymorpha* as a single circular form with no incorporation of foreign DNA. Biosci. Biotech. Biochem. **56**, 132–135.
37) Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N. *et al.* (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. A primitive form of plant mitochondrial genome. J. Mol. Biol. **223**, 1–7.
38) Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N. *et al.* (1992) Complete nucleotide sequence of the mitochondrial DNA from a livewort *M. polymorpha*. Plant Mol. Biol. Rep. **10**, 105–163.
39) Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N. *et al.* (1992) Transfer RNA genes in the mitochondrial genome from a liverwort, *Marchantia polymorpha*: the absence of chloroplast-like tRNAs. Nucl. Acids Res. **20**, 3773–3777.
40) Akashi, K., Sakurai, K., Hirayama, J., Fukuzawa, H. and Ohyama, K. (1996) Occurrence of nuclear-encoded tRNA Ile in mitochondria of the liverwort *Marchantia polymorpha*. Curr. Genet. **30**, 181–185.
41) Michael, F. and Dujon, B. (1983) Conservation of RNA secondary structures in two intron families including mitochondrial-, chloroplast- and nuclear-encoded members. EMBO J. **2**, 33–38.
42) Cech, T. R. (1988) Conserved sequences and structures of group I introns: Building an active site for RNA catalysis. Gene **73**, 259–271.
43) Rochaix, J. D. and Malnoe, P. (1978) Anatomy of the chloroplast ribosomal DNA of *Chlamydomonas reinhardtii*. Cell **15**, 661–670.
44) Fukuzawa, H., Kohchi, T., Shirai, H., Ohyama, K., Umesono, K., Inokuchi, H. *et al.* (1986) Coding sequences for chloroplast ribosomal protein S12 from the liverwort, *Marchantia polymorpha*, are separated far apart on the different DNA strands. FEBS Lett. **198**, 11–15.
45) Kuck, U., Choquet, Y., Schneider, M., Dron, M. and Bennoun, P. (1987) Structural and transcription analysis of two homologous genes for the P700 chlorophyll *a*-apoproteins in *Chlamydomonas reinhardtii*: evidence for *in vivo trans*-splicing. EMBO J. **6**, 2185–2195.
46) Fisher, S. E. J., Butler, M. D., Pan, Q. and Ruvkun, G. (2008) *Trans*-splicing in *C. elegans* generates the negative RNAi regulator ERI-6/7. Nature **455**, 491–496.
47) Kohchi, T., Ogura, Y., Umesono, K., Yamada, Y., Komano, T., Ozeki, H. *et al.* (1988) Ordered processing and splicing in a polycistronic transcript in liverwort chloroplasts. Curr. Genet. **14**, 147–154.
48) Ohyama, K., Kohchi, T., Ogura, Y., Oda, K., Yamato, K., Sano, T. *et al.* (1990) Gene organization and expression of chloroplast genome from a liverwort, *Marchantia polymorpha*. Bot. Mag. (Tokyo) **2**, 145–158.
49) Michel, F. and Westhof, E. (1990) Modeling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. J. Mol. Biol. **216**, 585–610.
50) Ohyama, K., Oda, K., Ohta, E. and Takemura, M. (1993) Gene organization and evolution of introns of a liverwort, *Marchantia polymorpha* mitochondrial genome. *In* Plant Mitochondria. (eds. Brennicke, A. and Kuch, U.). VCH Verlagsellshaft, VCH Publishers, Weinheim, pp. 115–129.
51) Ohyama, K. and Takemura, M. (2008) Molecular evolution of mitochondrial introns in the liverwort *Marchantia polymorpha*. Proc. Jpn. Acad., Ser. B **84**, 17–23.
52) Michel, F. and Lang, B. F. (1985) Mitochondrial class II introns encode proteins related to the reverse transcriptases of retroviruses. Nature **316**, 641–643.
53) Shaw, J. M., Feagin, K. J., Stuart, M. and Simpson, L. (1988) Editing of kinetopastid mitochondrial mRNAs by uridine addition and deletion generates converted amino acid sequence and AUG initiation codons. Cell **53**, 401–411.
54) Gualberto, J. M., Lamattina, L., Bonnard, G., Weil, J.-H. and Grienenberger, J.-M. (1989) RNA editing in wheat mitochondria results in the conservation of protein sequences. Nature **341**, 660–662.
55) Hoch, B., Maier, R. M., Appel, K., Igloi, G. L. and Kossel, H. (1991) Editing of a chloroplast mRNA by creation of an initiation codon. Nature **353**, 178–180.
56) Ohyama, K., Oda, K., Yamato, K., Ohta, E., Takemura, M. and Akashi, K. (1995) The mitochondrial genome of a liverwort, *Marchantia polymorpha*. *In* The Molecular Biology of Plant Mitochondria (eds. LecingsIII, C. S. and Vasil I. K.). Kluwer Academic Publishers, Dordrecht, pp. 597–633.
57) Ohyama, K., Ogura, Y., Oda, K., Yamato, K., Ohta, E., Nakamura, Y. *et al.* (1991) *In* Evolution of Life (eds. Osawa, S. and Honjo, T.). Springer-Verlag, Tokyo, pp. 187–198.
58) Fox, T. D. and Leaver, C. J. (1981) The *Zea mays*

mitochondrial gene coding cytochrome oxidase subunit II has an intervening sequence and does not contain TGA codons. Cell **26**, 315–323.
59) Covello, P.S. and Gray, M.W. (1989) RNA editing in plant mitochondria. Nature **341**, 662–666.
60) Rudinger, M., Polsakiewicz, M. and Knoop, V. (2008) Organellar RNA editing and plant-specific extensions of pentatricopeptide repeat proteins in Jungermannid but not in Marchantiid liverworts. Mol. Biol. Evol. **25**, 1–10.
61) Ozeki, H., Ohyama, K., Inokuchi, H., Fukuzawa, H., Kohchi, T., Sano, T. *et al.* (1987) Genetic system of chloroplasts. Cold Spring Harbor Symposium on Quantitative Biology **52**, 791–804.
62) Ozeki, H., Umesono, K., Inokuchi, H., Kohchi, T. and Ohyama, K. (1989) The chloroplast genome of plants: A unique origin. Genome **31**, 169–174.
63) Cavalier-Smith, T. (1987) The simultaneous symbiotic origin of mitochondria, chloroplasts and microbodies. Ann. New York Acad. Sci. **503**, 55–71.
64) Stern, D. B. and Lonsdale, D. M. (1982) Mitochondrial and chloroplast genomes of maize have a 12-kilobase DNA sequence in common. Nature **299**, 698–702.
65) Cho, Y., Qiu, T., Kuhlman, P. and Palmer, J. D. (1998) Explosive invasion of plant mitochondria by a group I intron. Proc. Natl. Acad. Sci. USA **95**, 14244–14249.
66) Sinclair, A. H., Berta, P., Palmer, M. S., Hawkins, J. R., Griffiths, B. L., Smith, M. J. *et al.* (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature **346**, 240–244.
67) Koopman, P., Munsterberg, A., Capel, B., Vivian, N. and Lovell-Badge, R. (1990) Expression of a candidate sex-determining gene during mouse testis differentiation. Nature **348**, 450–452.
68) Ross, M. T., Grafham, D. V., Coffey, A. J., Scherer, S., McLay, K., Muzny, D. *et al.* (2005) The DNA sequence of the human X chromosome. Nature **434**, 325–337.
69) Charlesworth, B. and Charlesworth, D. (2000) The degeneration of Y chromosomes. Philos. Trans. R. Soc. Lond. B. Biol. Sci. **355**, 1563–1572.
70) Charlesworth, D. (2002) Plant sex determination and sex chromosomes. Heredity **88**, 94–101.
71) Westergaard, M. (1958) The mechanism of sex determination in dioecious flowering plants. Adv. Genet. **9**, 217–281.
72) Liu, Z., Moore, P. H., Ma, H., Ackerman, C. M., Ragiba, M., Yu, Q. *et al.* (2004) A primitive Y chromosome in papaya marks incipient sex chromosome evolution. Nature **427**, 348–352.
73) Ainsworth, C., Parker, J. and Buchanan-Wollaston, V. (1998) Sex determination in plants. Curr. Top. Dev. Biol. **38**, 167–223.
74) Juarez, C. and Banks, J.A. (1998) Sex determination in plants. Curr. Opin. Plant Biol. **1**, 68–72.
75) Bull, J.J. (1983) Evolution of Sex Determining Mechanisms (Benjamin-Cummings, Menlo Park, CA).
76) Qiu, Y. L. and Palmer, J. D. (1999) Phylogeny of early land plants: insights from genes and genomes. Trends in Plant Sci. **4**, 26–30.
77) Newton, M. E. (1984) The Cytogenetics of Bryophytes. *In* The Experimental Biology of Bryophytes (eds. Dyer, A.F. and Duckett, J. G.). Academic Press, London, pp. 65–96.
78) Haupt, G. (1932) Beiträge zur Zytologie der Gattung *Marchantia* (L.). Z. Indukt. Abstamm. Vererbungsl. **62**, 367–428.
79) Lobeer, G. (1934) Die Zytologie der Lebermoose mit besonderer Berücksichtigung allgemeiner Chromosomenfragen. Jahrb. Wiss. Bot. **80**, 567–817.
80) Okada, S., Sone, T., Fujisawa, M., Nakayama, S., Takenaka, M., Ishizaki, K. *et al.* (2001) The Y chromosome in the liverwort *Marchantia polymorpha* has accumulated unique repeat sequences harboring a male-specific gene. Proc. Natl. Acad. Sci. USA **98**, 9454–9459.
81) Fujisawa, M., Hayashi, K., Nishio, T., Bando, T., Okada, S., Yamato, K. T. *et al.* (2001) Isolation of X- and Y-chromosome-specific DNA markers from a liverwort, *Marchantia polymorpha*, by representational difference analysis. Genetics **159**, 981–985.
82) Yamato, K. T., Ishizaki, K., Fujisawa, M., Okada, S., Nakayama, S., Fujishita, M. *et al.* (2007) Gene organization of the liverwort Y chromosome reveals distinct sex chromosome evolution in a haploid system. Proc. Natl. Acad. Sci. USA **104**, 6472–6477.
83) Ishizaki, K., Shimizu-Ueda, Y., Okada, S., Yamamoto, M., Fujisawa, M., Yamato, K. T. *et al.* (2002) Multicopy genes uniquely amplified in the Y chromosome-specific repeats of the liverwort *Marchantia polymorpha*. Nucleic Acids Res. **30**, 4675–4681.
84) Nakayama, S., Fujishita, M., Sone, T. and Ohyama, K. (2001) Additional locus of rDNA sequence specific to the X chromosome of the liverwort, *Marchantia polymorpha*. Chromosome Res. **9**, 469–473.
85) Fujisawa, M., Nakayama, S., Nishio, T., Fujishita, M., Hayashi, K., Ishizaki, K. *et al.* (2003) Evolution of ribosomal DNA unit on the X chromosome independent of autosomal units in the liverwort *Marchantia polymorpha*. Chromosome Res. **11**, 695–703.
86) Skaletsky, H., Kuroda-Kawaguchi, T., Minx, P. J., Cordum, H. S., Hillier, L., Brown, L. G. *et al.* (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature **423**, 825–837.
87) Charlesworth, B. (1978) Model for evolution of Y chromosomes and dosage compensation. Proc. Natl. Acad. Sci. USA **75**, 5618–5622.

(Received Dec. 12, 2008; accepted Jan. 22, 2009)

Profile

Kanji Ohyama was born in 1939, Kyoto, Japan. After he received his Doctor Degree of Agriculture from Kyoto University, he started his research career as Research Associate with Professor Armin C. Braun at The Rockefeller University on Plant Biology, and moved to Laboratory of Plant Biotechnology, National Research Council of Canada(NRCC), Saskatoon, Canada. After his post doctoral training, he joined Plant Research Group of NRCC as a staff and started investigation on DNA uptake using plant cells. This was the first demonstration of DNA uptake by plant protoplasts. In 1974 he began studies on the genome research using plant chroloplasts. In 1980 he joined to Laboratory of Biochemistry, Faculty of Agriculture, Kyoto University and continued his research career being established by series of research on plant genomes of chroloplasts, mitochondria and male sex chromosome from a liverwort *Marchantia polymorpha*. During his research the exceptionally discovery was the elucidation of trans-splicing systems in chroloplast genome. He was promoted to Professor of Faculty of Agriculture in 1990 and moved to Graduate School of Biostudies, Kyoto University in 1999 and served as the Head. He was awarded The Japan Bioscience, Biotechnology and Agrochemistry Society Award in 1995, and The Prize of Japan Academy in 2008. His current research interests are bioproduction of pharmaceutical compounds using plant species.