



Recent highlights of the China Rice Functional Genomics Program

Yongbiao Xue¹, Jiayang Li¹ and Zhihong Xu²

¹Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100080, China

²Peking University, Beijing 100871, China

To discover genes essential for rice genetic improvement, the Ministry of Science and Technology of China has started the China Rice Functional Genomics Program (CRFGP). During the past three years, the CRFGP has focused on developing rice functional genomics tools and resources to identify genes of agronomic significance. Here, we highlight recent progress made by the CRFGP and discuss the possible integration of global resources for rice functional genomics.

Rice (*Oryza sativa* L.) is the most important food crop for more than half the world's population. Because of its small genome size (~430 Mb) [1], ease of genetic manipulation and generic synteny with other cereal genomes, rice has emerged as a major model species in plant genome research. Recent progress in sequencing the rice genome has made rice a favorite plant for functional genomics research [2–5].

Genome sequencing has predicted that the rice genome has around 50 000 to 60 000 genes [2,3], which is nearly twice the number of genes predicted in *Arabidopsis* [6]. Interestingly, nearly half of the predicted rice genes have no obvious homologs in *Arabidopsis* although there is substantial homology with available expressed sequence tags (ESTs) in other cereals [2–5]. Two different strategies have been used for sequencing rice genomes. Two working drafts for rice subspecies *indica* and *japonica* genomes were produced by using a whole-genome shotgun strategy [2,3]. The *indica* sequence is freely available to the public [2] but the *japonica* sequence is available through agreement [3]. The International Rice Genome Sequencing Project (IRGSP) [7] announced in December 2002 that a high-quality draft sequence for the *japonica* genome had been produced, and made freely available to the public. In fact, less than 35% of the annotated genes from the finished sequences of rice chromosomes 1 and 4 have functional notes [4,5]. Thus, rice functional genomics will greatly help to assign the function of most rice genes.

Unlike the organized consortium for rice genome sequencing [7], there are as yet no substantial international collaborative efforts on rice functional genomics. However, several countries, including Japan, Korea and China, have made a strong commitment to invest in rice functional genomics. In China, in addition to supporting rice genome sequencing, the Ministry of Science and

Technology has funded the China Rice Functional Genomics Program (CRFGP) since 1999 under the National Basic Sciences Initiatives [8]. Twenty research groups are engaged in the program, drawn from the Institutes of Genetics and Developmental Biology, Botany and Plant Physiology, and Ecology of the Chinese Academy of Sciences and Peking University, Tsinghua University, South China Agricultural University, Yangzhou University, Wuhan University and Central China Agricultural University. The mission of the CRFGP is to discover and make use of genes to improve agronomic performance of rice as well as other crops. In the past three years, the CRFGP has focused on developing tools and resources for functional genomics and identifying genes essential for rice breeding and genetic improvement. Here we highlight some of these advances and also discuss the possible integration of global resources for rice functional genomics research.

Development of rice functional genomics tools

Functional genomics is characterized by high-throughput gene functional annotation. Research tools and resources, including genome-wide mutant generation, transcript profiling and data mining, are being developed in several plant species for functional genomics studies [9–11]. In the past three years, the CRFGP has focused on developing similar functional genomics tools for rice. These tools include the creation and collection of mutants by radiation- or chemical-induced mutagenesis, and insertional mutagenesis – using transfer DNA (T-DNA) of *Agrobacterium* and maize *activator/dissociation* (*Ac/Ds*) transposons; cDNA microarrays; transformation-competent artificial chromosome (TAC)-based gene cloning strategies; and bioinformatics. Furthermore, we are also identifying genes of agronomic significance by map- and homology-based cloning approaches.

The exploitation of genetic variation forms a basis for assigning the function of genome sequences. In plants, two strategies have been widely used to generate genetic variation: chemical- or radiation-induced mutagenesis and insertional mutagenesis using T-DNA or transposable elements, for example, *Ac/Ds* (Box 1). Several T-DNA and transposon-tagged populations of rice have been generated [12–15]. The CRFGP is using both T-DNA and *Ac/Ds* as well as radiation to produce rice mutants. To create a mutant population, a construct containing *Ac/Ds* in conjunction with an enhancer trapping strategy has

Corresponding author: Yongbiao Xue (ybxue@genetics.ac.cn).

Box 1. Insertional mutagenesis in rice

Agrobacterium T-DNA (transfer DNA) construct with transposable elements (*Ac/Ds*) from maize

Agrobacterium-mediated transformation results in the integration of the T-DNA in the plant nuclear genome (Gene 1) (Fig. I). A large-scale transformation experiment will produce a T-DNA tagging population for functional genomics in rice. *Ac/Ds* within the T-DNA could transpose into another gene either autonomously (*Ac*) or nonautonomously (*Ds*) helped by *Ac* in *trans*. LB, left border; RB, right border; *Hgr^R*, gene conferring hygromycin resistance.

Agrobacterium-mediated transformation

Immature or mature embryos collected from rice grains are cocultivated with *Agrobacterium* harboring a T-DNA construct (Fig. II). Calli are subsequently induced on a selective medium. Under appropriate conditions, transformed somatic embryos are formed and eventually lead to the production of transgenic seedlings.

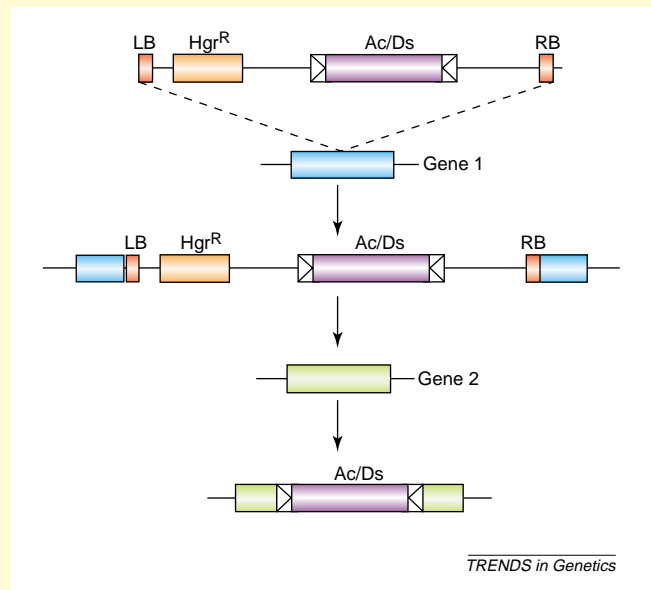


Fig. I.

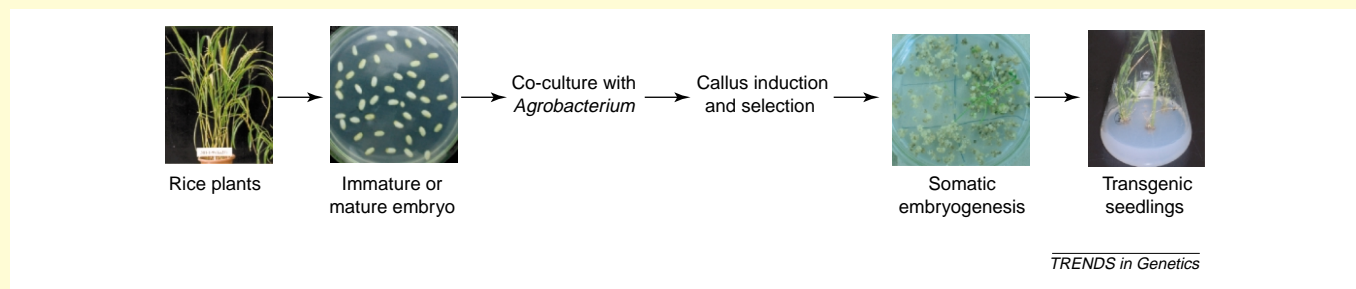


Fig. II.

been used to transform rice (*O. sativa* ssp. *japonica* var. Zhonghua 11 and var. Nipponbare). So far, more than 5000 *Ac/Ds* and 20 000 independent T-DNA insertion lines have been obtained, and several thousand insertion sites mapped on the rice chromosomes. Insertional mutants affecting several aspects of rice growth and development related to important agronomic traits are being characterized, and several genes affected have been cloned (J. Zhang, pers. commun.; Y. Xue, unpublished). In addition, more than several thousand mutant lines induced by radiation have been obtained and a spectrum of interesting mutant phenotypes observed. These mutants, together with those already available, will serve as an indispensable resource for genetic discovery in rice.

Expression profiling

More than 110 000 ESTs from several cDNA libraries constructed from rice (*O. sativa* ssp. *indica* GLA4) tissues under normal and stressed conditions have been sequenced (<http://ncgr.ac.cn/EST.html/>). Taking advantage of these EST collections, a cDNA chip containing >12 000 unique ESTs has been used to determine the gene expression profiles during normal growth and

development, as well as under stress conditions such as drought and salinity. More than 100 genes have been found to be either induced or repressed during rice pollination, and altered gene expression patterns have been confirmed by quantitative PCR analysis. In addition, hormone (auxin, brassinosteroid and abscisic acid)-regulated genes are also being profiled and analyzed. Functional annotation has begun to shed light on the molecular mechanisms of rice pollination and hormonal responses, thus providing useful resources, including promoters, for future genetic manipulations (Y. Xue, unpublished). Recently, we have also initiated an 'ORFeome' project to study protein functions and interactions in rice, including cloning of full-length open reading frames (ORFs) into plant and yeast expression vectors for over- and under-expression in plants and yeast two-hybrid screening.

Map-based gene cloning has been successfully used in rice [16,17], although it is more tedious and time-consuming because of the larger genome size and longer life cycle of rice compared to *Arabidopsis*. To make better use of genome sequence information, we take advantage of binary bacterial artificial chromosome (BIBAC) and TAC vector systems, which are effective for transferring large

DNA fragments into plant genomes by *Agrobacterium*-mediated transformation [18,19]. So far, several TAC libraries have been constructed from rice genomic DNA, and it has been shown that large fragments of genomic DNA from TAC clones could be efficiently integrated into the rice genome [20]. Currently, we are sequencing TAC ends to align them on to the finished rice genome map to generate a TAC physical map. Eventually, the clones for a TAC contig covering a target gene can be used to transform and complement mutants and for other genetic manipulations. This will lead to the identification of the target gene without using a large segregating population, which is required for fine-scale mapping in map-based cloning. Using this strategy, several genes affecting fertility have already been cloned (Y. Liu, pers. commun.).

Bioinformatics

Clearly, making sense of the vast amount of gene information generated requires the techniques of bioinformatics. The CRFGP has developed an integrated database for rice gene information generated by the Program itself and by other related projects worldwide. Effective bioinformatics tools for data analysis and management are also being developed. Currently, the database contains information on rice insertional mutants, TAC end sequences and ESTs. The results from the microarray analysis are also being integrated. The database can be accessed at the Program's website (<http://www.rifgp.ac.cn>).

Discovery of genes of agronomic significance

In recent years, hybrid rice breeding has been a driving force for the substantial increase in rice grain yield seen during the past 20 years in China (Box 2). In hybrid rice production, one limitation is related to the narrow genetic background of fertility restorer lines used, which reduces the possibility of selecting optimal parental combinations for making hybrids. However, it is very tedious and time-consuming to introduce a fertility restorer gene into an elite alien genetic background by conventional breeding. This obstacle could be overcome by specifically introducing fertility restorer genes into alien elite lines through a transgenic approach. Recently, a new type of rice fertility restorer gene, *Rf3*, for the wild-abortive type (WA) of CMS was identified (Y. Liu, pers. commun.), which will make this molecular approach feasible. Manipulation of elite lines with the *Rf1* gene could widen the donor lines used in hybrid rice breeding, thus potentially improving the grain quality of most hybrid rice lines. Candidates for several other genes important for hybrid rice breeding, for example, hybrid sterility genes *Sa* and *Sb* and temperature-sensitive genic male sterility (TGMS) genes, have also been identified. By using these genes we could further explore the subspecies hybrid vigor and enhance the breeding efficiency of hybrid rice.

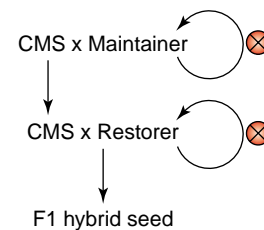
Poor grain quality has sometimes been a problem associated with hybrid rice, so breeding high-yielding rice cultivars with good grain quality is still highly desirable. Thus, the discovery of genes affecting plant growth and development traits, such as tillering, fertility and plant architecture, offers the potential for further genetic improvement of rice. The CRFGP has specifically targeted and cloned several such genes by a map-based gene cloning strategy.

Box 2. Schemes for hybrid rice production

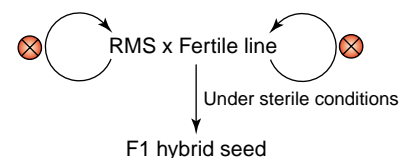
In three-line hybrid rice production (Fig. IIIa), cytoplasmic male sterile (CMS), fertility restorer and maintainer lines are used. To make a hybrid, a CMS line is crossed with a fertility restorer line. CMS is maintained by intercrossing CMS and maintainer lines. Although this scheme is complicated compared to two-line hybrid rice production, it is still widely used for hybrid rice production; for example, an elite *indica* hybrid Shanyou 63 produced this way is widely cultivated in China. The maintainer and restorer lines are inbred lines.

In two-line hybrid rice production (Fig. IIIb), a revertible male sterile (RMS) line and a fertile line are used for making hybrids. The fertility of the RMS line can be influenced by environmental factors, such as photoperiod and/or temperature. Under one condition, for example lower temperature, the so-called temperature-sensitive genic male sterile (TGMS) line is sterile and thus suitable for making hybrids. Under another condition, for example, higher temperature, the TGMS line is fertile for self-renewal. Several elite hybrid rice lines have been produced in this way and are widely grown in China. Both lines used are inbred lines.

(a) Three-line hybrid rice production



(b) Two-line hybrid rice production



TRENDS in Genetics

Fig. III.

Tillering, where multiple stems with seed ears are produced from a single plant, is an important agronomic trait related to yield. A rice gene called *Monoculm1* (*MOC1*) has been cloned recently [21]. The *moc1* mutant plants (Fig. 1) have only a main culm without any tillers due to a defect in the formation of tiller buds. *MOC1* encodes a putative GRAS family nuclear protein, which is mainly expressed in the axillary buds and functions to initiate axillary buds and to promote their outgrowth. Rice with a mutation in *Brittle Culm1* (*BC1*) has been used to breed rice cultivars whose culms can be used for animal feed owing to a reduced cellulose content compared with conventionally cultivated rice varieties (Q. Qian, pers. commun.). Recently, the *BC1* gene, which is involved in cellulose biosynthesis, has been cloned by a map-based strategy (J. Li, unpublished). Thus, it is possible to regulate the expression of *BC1* or its orthologs in rice and other cereals to alter the cellulose content as desired.

Apart from the map-based cloning of genes of agronomic significance, we are also interested in transcription



Fig. 1. Rice plants with a mutation at *Monoculm1* (*moc1*) and normal (wild-type) tiller (WT).

factors. Transcription factors are key regulators for plant growth and development as well as physiological and biochemical processes. We have focused our study on a genome-wide identification and functional analysis of transcription factors, particularly those involved in controlling grain quality and yield in rice. At present, expression profiling analysis of ~400 rice transcription factors using a cDNA chip has revealed various families of transcription factors involved in different developmental processes and cellular responses to environmental factors (H-W. Xue, pers. commun.). Transcription factors or elements with gene activation function during rice grain formation are being isolated through a transactivation system based on a modified yeast two-hybrid system. In principle, a cDNA fused with a GAL4-binding domain could encode a transcription factor if the fusion is capable of inactivating a reporter gene expression in yeast. In addition, candidates including members of MADS, AP2, bZIP, MYB and HB classes of transcription factors involved in these processes, have been identified and functional characterization of these genes is being carried out by various reverse genetic approaches, from which we expect to identify spectra of phenotypes useful for rice breeding and genetic improvement.

Application of functional genomics knowledge in rice breeding

The success of traditional rice breeding has relied on the use of genes affecting plant height (shorter plants put more photosynthate into grain) and disease resistance as well as the exploitation of hybrid vigor. It is reasonable to believe that genes discovered through functional genomics will greatly improve rice breeding.

Through comparative analysis, genome-wide intra-specific sequence variations, including insertions and deletions as well as single nucleotide polymorphisms, have been identified [3,4,22]. These variations are probably associated with phenotypic differences, as is seen between the two major rice subspecies, *indica* and

japonica. Importantly, these genome-wide sequence variations will provide an unprecedented opportunity to develop markers for molecular-assisted breeding in rice.

An immediate application of discoveries made by rice functional genomics is for genetic improvement of rice by a transgenic approach. To build up a pipeline from genes to traits, we have developed a platform for genetic improvements of key agronomic traits including a robust transformation technology and controllable expression systems. This can be illustrated by using an antisense *Waxy* gene and its promoter, to produce stable transgenic lines with a range of reduced amylose contents. This trait has been incorporated into elite fertility restorer lines as well as cultivars commonly grown in China [23]. In addition, reliable transformation procedures have been developed for most recalcitrant cultivated *indica* rice lines in China and the generated transgenic lines have been shown to be stable over several generations in field tests [23].

In some cases, several genes have to be transferred together to generate a trait of interest (e.g. if they encode several components of a biosynthetic pathway). Therefore, it is highly desirable to have a multiple gene transfer system for genetic engineering in rice. To be able to transfer multiple genes simultaneously, a shuttle binary vector system has been developed, which allows the assembly of several target genes into a single vector. By using this system, up to nine genes have been assembled and stably transferred into rice [24]. Currently, several genes discovered in the Program are being tested with the aim of improving rice yield and grain quality using this strategy.

Integration of global resources for rice functional genomics

Recently, several attempts have been made to integrate global resources for rice functional genomics research. In particular, the International Rice Research Institute (IRRI) has been a strong advocate for an international functional genomics consortium [25]. In Japan and Korea, major efforts have been invested in developing rice functional genomics tools and identifying genes of agronomic significance. Other countries, such as the USA, Australia, France and The Netherlands, also have programs on rice functional genomics. Although the researchers involved in these programs do communicate from time to time, it is apparent that a more coordinated effort is required to make better and more effective use of existing resources within rice functional genomics. Needless to say, such an effort could have an immense impact on the future security of world food. By joining forces, the pace of functional gene discovery and genetic manipulations in rice could be increased. Furthermore, the recent completion of the rice genome sequence has provided an opportunity to coordinate an international program for rice functional genomics. We believe that this program should be modeled on other similar international initiatives of functional genomics. Importantly, a giving-and-sharing principle should be the foundation for such an international initiative. As part of our effort towards this goal, we are going to organize the First International Symposium of Rice Functional Genomics, to be held in Shanghai in late 2003.

Conclusion

Rice has become a model species for plant functional genomics research. In a sense, rice is unique because it is a model plant as well as a staple crop. By integrating global resources, an international program could be put in place with clear and achievable targets. The CRFGP is eager to contribute and share the rice functional genomics tools and resources it has developed with such an international endeavor. Only by working together can the benefits of rice functional genomics be fully realized at the earliest opportunity, in terms of improved crops of rice and other species.

Acknowledgements

The CRFGP is funded by the Ministry of Science and Technology of China (G1999011600). We are grateful to the Chinese Academy of Sciences for supporting rice functional genomics, and, in particular, to all the scientists involved in the Program. We thank Jingliu Zhang, Da Luo, Yaoguang Liu, Hongwei Xue, Kang Chong, Qian Qian, Hai Huang and Guohua Liang for communicating their unpublished results. We also thank Bin Han for sequencing ESTs, and Jianru Zuo and Weicai Yang for critically reading the manuscript.

References

- Arumuganathan, K. and Earle, E.D. (1991) Nuclear DNA content of some important plant species. *Plant Mol. Biol. Rep.* 9, 208–218
- Goff, S.A. *et al.* (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296, 92–100
- Yu, J. *et al.* (2002) A draft sequence of the rice genome (*Oryza sativa* ssp. *indica*). *Science* 296, 79–91
- Feng, Q. *et al.* (2002) Sequence and analysis of rice chromosome 4. *Nature* 420, 316–321
- Sasaki, T. *et al.* (2002) The genome sequence and structure of rice chromosome 1. *Nature* 420, 312–316
- The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815
- Sasaki, T. and Burr, B. (2000) International rice genome sequencing project: the effort to completely sequence the rice genome. *Curr. Opin. Plant Biol.* 3, 138–141
- Xue, Y. and Xu, Z. (2002) An introduction to the China rice functional genomics program. *Comp. Funct. Genomics* 3, 161–163
- Sessions, A. *et al.* (2002) A high-throughput *Arabidopsis* reverse genetics system. *Plant Cell* 14, 2985–2994
- Boyes, D.C. *et al.* (2001) Growth stage-based phenotypic analysis of *Arabidopsis*: a model for high throughput functional genomics in plants. *Plant Cell* 13, 1499–1510
- Francki, M. and Appels, R. (2002) Wheat functional genomics and engineering crop improvement. *Genome Biol.* 3, reviews1013.1–1013.5
- Chin, H.G. *et al.* (1999) Molecular analysis of rice plants harboring an Ac/Ds transposable element-mediated gene trapping system. *Plant J.* 19, 615–623
- Greco, R. *et al.* (2001) Early and multiple Ac transposition in rice suitable for efficient insertional mutagenesis. *Plant Mol. Biol.* 46, 215–227; Erratum 46, 763–764
- Jeon, J.S. *et al.* (2000) T-DNA insertional mutagenesis for functional genomics in rice. *Plant J.* 22, 561–570
- Kohli, A. *et al.* (2001) Tagged transcriptome display (TTD) in *indica* rice using Ac transposition. *Mol. Genet. Genomics* 266, 1–11
- Yano, M. *et al.* (2000) *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. *Plant Cell* 12, 2473–2483
- Yamanouchi, U. *et al.* (2002) A rice spotted leaf gene, *Sp11*, encodes a heat stress transcriptional factor protein. *Proc. Natl. Acad. Sci. U. S. A.* 99, 7530–7535
- Hamilton, C.M. *et al.* (1996) Stable transformation of intact high molecular weight DNA into plant chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* 93, 9975–9979
- Liu, Y.G. *et al.* (1999) Complementation of plant mutants with large genomic DNA fragments by a transformation-competent artificial chromosome vector accelerates positional cloning. *Proc. Natl. Acad. Sci. U. S. A.* 96, 6535–6540
- Liu, Y.G. *et al.* (2002) Development of new transformation-competent artificial chromosome vectors and rice genomic libraries for efficient gene cloning. *Gene* 282, 247–255
- Li, X. *et al.* (2003) Control of tillering in rice. *Nature* 422, 618–621
- Han, B. and Xue, Y. (2003) Genome-wide intraspecific DNA-sequence variations in rice. *Curr. Opin. Plant Biol.* 6, 134–138
- Liu, Q. *et al.* (2003) Stable inheritance of the antisense *waxy* gene in transgenic rice with reduced amylose level and improved quality. *Transgenic Res.* 12, 71–82
- Lin, L. *et al.* Efficient linking and transfer of multiple genes by a multi-gene assembly and transformation vector system. *Proc. Natl. Acad. Sci. U. S. A.* 100, 5962–5967
- Fischer, K.S. *et al.* (2000) Collaborations in rice. *Science* 290, 279–280

Genetics of Disease Current Opinion in Genetics & Development

The current issue of *Current Opinion in Genetics & Development* focuses on the genetics of disease and the tremendous insights provided into normal human biology through the study of disease phenotypes, including:

The use of telomere probes to investigate submicroscopic rearrangements associated with mental retardation
Jonathan Flint and Samantha Knight

Genetics and genomics of behavioral and psychiatric disorders
Ken Inoue and James R Lupski

Modifier genes and protective alleles in humans and mice
Joseph H. Nadeau

Finding genes underlying risk of complex disease by linkage disequilibrium mapping
Andrew G. Clark

Rodent models of genetic disease
Roger D. Cox and Steve D.M. Brown

Muscular dystrophies: genes to pathogenesis
Isin Dalkilic and Louis M. Kunkel

Transcriptional derepression as a cause of genetic diseases
Davide Gabellini, Rossella Tupler and Michael R. Green

Chromatin remodeling and human disease
Cheng Huang, Emily A. Sloan and Cornelius F. Boerkoel

Nuclear genes in mitochondrial disorders
Massimo Zeviani, Antonella Spinazzola and Valerio Carelli