Proceedings of ANAS (Biological Sciences), 65(5-6): 140-144 (2010)

Gene Discovery and Advances in Biotechnology

Shakhira M. Zakhrabekova[1]*, Leif Lundh[2], Mats Hansson[1]
[1]*Department of Biochemistry, Lund University, Box 124, SE-22100 Lund, Sweden*
[2]*Department of Cell and Molecular Biology, Gothenburg University, Box 462, 405 30 Gothenburg, Sweden*

History of biotechnology is very old and covers such well known and important applications as brewing beer, fermenting vine, baking bread, producing of monoclonal antibodies and many others. Discovery of recombinant DNA techniques culminated with the birth of genetic engineering and became a scientific beginning in the epoch of modern biotehnology. Today's advances in biotechnology could not be achieved without the major sources for gene discovery: availablity of complete genomic sequence information for the organism, access to high-throughput genotyping technologies, success in developing computational biology tools and developing of different molecular markers techniques. There are four major industrial applications in modern biotechnology: crop production and agriculture; non-food uses of crops and other products; medicine (health care) and environmental uses. In this article, we will in more details describe advantages and disadvantages of several molecular marker techniques and make a short review of the different applications in modern biotechnology.

Keywords: *biotechnology, genetic variation, genetic markers, crop production, environment, medicine*

Genomic resources and tools for gene discovery

DNA markers, quantitative trait loci (QTL), and marker-assisted selection (MAS)

All organisms are subjected to mutations, leading to a natural genetic variation (polymorphism). For these variations to be valuable to researchers, it is necessary that they are heritable and sensible. Natural variations must be recognizable by phenotype or have a genetic mutation distinguishable through molecular techniques.

Different types of genetic variation at the DNA level can be classified in four classes: SNPs - single nucleotide polymorphism, indels - insertions or deletions of nucleotide sequences, inversion and rearrangements (Table 1).

DNA marker technology can be used to detect these mutations, discover the natural genetic variations and facilitate the construction of high-resolution genetic linkage maps for different species. More recent and popular types of genetic markers include: restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), randomly amplified polymorphic DNA (RAPD), single nucleotide polymorphism (SNP), microsatellite, allozymes, mitochondrial DNA and expressed sequence tag (EST) markers. There are two types of molecular markers: type I represent markers, which were identified during analysis of known genes and class II are associated with genes of unknown functions (O'Brien, 1991; Table 2). The power of the various marker types can be defined by their PIC (polymorphic information content) value (Botstein et al., 1980). PIC value of a marker detects polymorphism in a population and can give an idea of the power of the usefulness of molecular marker.

Characteristics of molecular markers and their potential applications

RFLP-markers

The traditionall technique for detecting RFLPs involves four steps: restriction of genomic DNA with restriction endonucleases; separation of the resulting DNA fragments by length in agarose gel electrophoresis, transfer of procedure; hybridization of the membrane to a labeled DNA probe and determination if the size of detected fragments varies between individuals (Botstein et al., 1980). Most recent analyses exchange the laborious Southern blot method with techniques based on the PCR. If flanking sequences are known for a locus, the piece of DNA, containing the RFLP region is amplified via PCR. If the length polymorphism is caused by a relatively large indels, gel electrophoresis of the PCR products should directly detect the size difference. However, if the length polymorphism is caused by SNP mutation at a restriction site, PCR products must be digested with a restriction enzyme to reveal the RFLP.

The PIC value of RFLP markers in revealing

*E-mail: zakhrabekova@yahoo.se

Table 1. Different types of genetic variation and their characteristics

Genetic variation	Characteristics
SNPs	Nucleotide base substitutions
Indels	Insertions or deletions of nucleotide sequences within a locus
Inversion	Inversion of a segment of DNA within a locus
Rearrangement	Rearrangement of DNA segments around a locus of interest

Table 2. Major types of DNA markers and their applications

Marker name	Type	The usefulness of molecular markers can be measured based on their PIC value	Applications
RFLP	I	relatively low	Linkage mapping
RAPD	II	intermediate	Fingerprinting for population studies
AFLP	II	high	Linkage mapping, population studies
Microsatellite (SSR)	Normally Type II	high	Linkage mapping, population studies
EST	I	low	Linkage mapping, physical mapping, comparative mapping
SNP	I or II	high	Linkage mapping

genetic variation is relatively low compared to more recently developed markers and techniques discussed below. Insertions or deletions of nucleotide sequences and rearrangements of regions containing restriction sites are possibly widespread in the genomes of most species, but the chances of such occasion within the locus of interest should be rare. The main power of RFLP markers is that they are codominant markers, i.e., both alleles in an individual are detected in the analysis.

RAPD-markers

RAPD procedures (Welsh and McClelland, 1990; Williams et al., 1990) use PCR to randomly amplify unknown fragments of nuclear DNA with an identical pair of primers 8-10 bp in length. Because of the low annealing temperatures and the short length of the primers, the probability of amplifying multiple products is very high, with each product likely representing a different locus. RAPD polymorphisms can occur due to the SNP at the primer binding sites and due to insertion or deletion in the regions between the sites. The PIC values of RAPD markers are relatively high, but RAPDs may not be as informative as AFLPs, because fewer loci are generated simultaneously.

AFLP - markers

AFLP is a PCR-based fingerprinting method (Vos et al., 1995) that combines the power of both RFLP and RAPD methods. While RFLP allows analysis of one locus at a time, AFLP has the potential to analyze number of loci simultaneously. The strength and reproducibility of AFLP analysis is extremely high. One major weakness of AFLP method is the need for special equipment such as automated gene sequencers for electrophoretic analysis of fluorescent labels.

Microsatellites

Microsatellites are composed of multiple copies of tandemly arranged simple sequence repeats (SSRs), which are more or less evenly spread in the genome on all chromosomes. They have been detected inside gene coding regions (Liu et al., 2001), introns, and in the non-gene sequences.

Most microsatellite loci are relatively small and can be amplified using PCR. The abundant amount of alleles per locus results in very high PIC values for microsatellite markers. Even genomic distribution and high PIC number makes microsatellite markers very popular and useful in population studies, although application of microsatellite markers requires a large amount of effort. Each microsatellite locus has to be determined and PCR primers designed in its flanking region.

Single nucleotide polymorphism

Single nucleotide polymorphism is caused by point mutations. SNPs are the most abundant in any organism and usually restricted to one of two alleles, although in theory, a SNP within a locus can produce up to four alleles, each carrying one of four bases at the SNP site: A, T, C, and G. SNPs are evolutionarily conserved and have been advised as markers for quantitative trait loci analysis instead of microsatellites. There are many different methods available for SNP genotyping, among which microarray (gene

chip) technology and quantitative PCR are very useful in medical and clinical studies (Hacia et al., 1999).

Crop production and agriculture

Improving grain yield, resistance of crops to environmental stresses, improving taste, texture and appearance of food, increasing nutritional qualities of food crops, production of novel substances in crop plants are very important applications of modern agricultural biotechnology (Figure 1).

The limitations on the commercial use of genetically modified plants, has led to a rise of interest in investigating natural biodiversity to increase the productivity, quality and nutritional value of crops. The exploration of such natural variations is very actual both in crop species and in the model species *Arabidopsis thaliana*.

Number of studies, concerning the genetics of crop yield, has been carried out in recent years, including a broad class of species like rice, barley, soybean and tomato (Xiao et al., 1996; Concibido et al., 2003; Huang et al., 2003; Septiningsih et al., 2003; Gur and Zamir, 2004). A recent study in rice discovered the QTL Gn1a, which improve crop yield

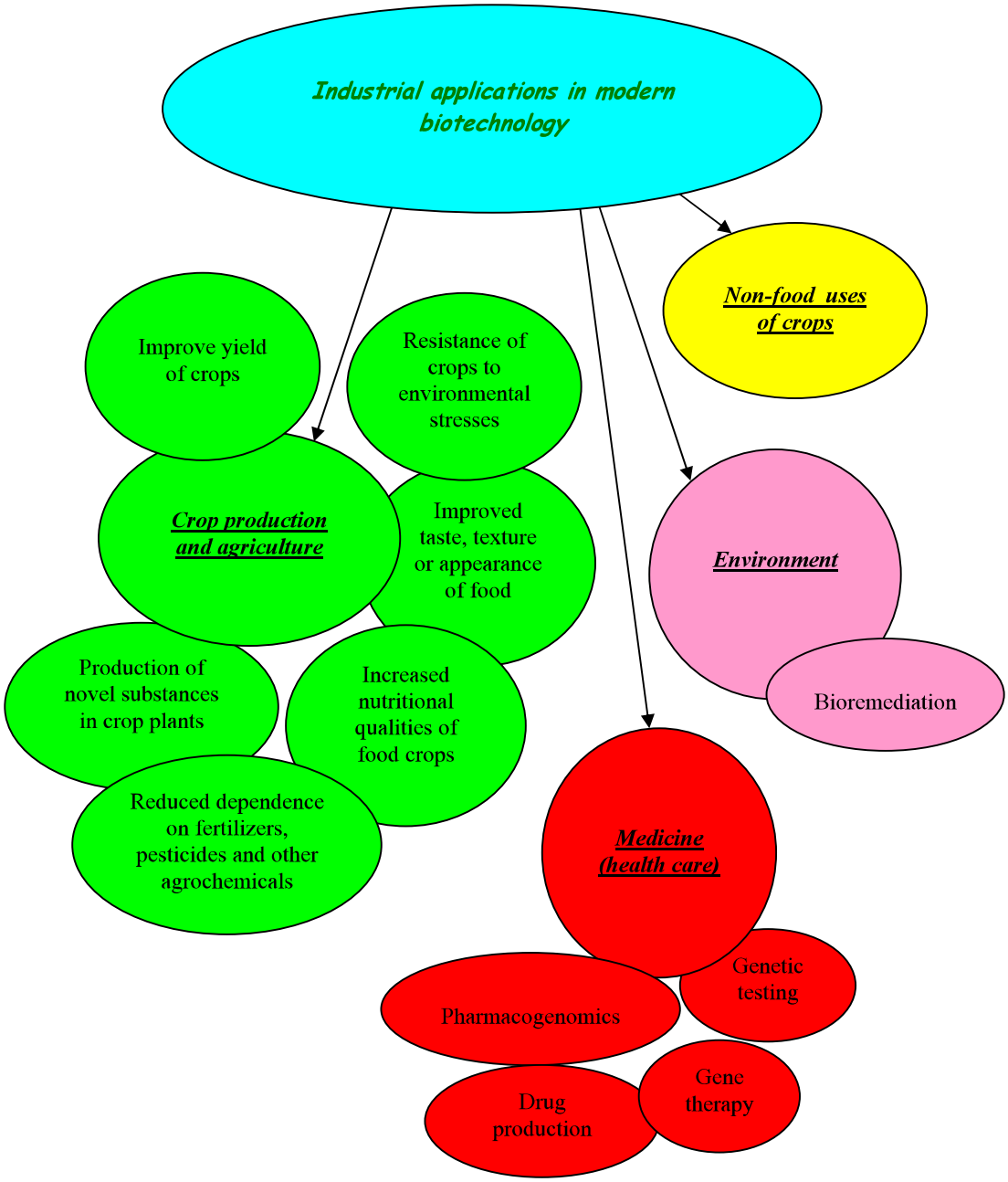

Figure 1. Industrial applications in modern biotechnology.

and encodes a cytokinin oxidase/dehydrogenase (OsCKX2), an enzyme that break down the phytohormone cytokinin. Down regulation of OsCKX2 lead to the accumulation of cytokinin in inflorescence meristems, resulting in an elevated number of reproductive organs and improved grain yield (Ashikari, 2005). Improving of crop yield can be as well achieved by so called QTL pyramiding. Studies, performed on rice showed that combined loci for grain number and plant height in the same genetic background generated lines that exhibited both beneficial traits (Ashikari, 2005).

Increasing nutritional qualities, such as protein, starch and oil content of food crops are very much in the focus of modern biotechnology. It was shown that, introgression of a high-grain-protein quality trait from *Triticum dicoccoides* into in the Canadian durum wheat (*Triticum turgidum* L.) resulted in elevated protein quality lines of durum wheat and markedly improved the quality of pasta made from flour of wheat carrying the QTL (Kovacs et al., 1998).

Vitamins, pigments and antioxidants have been studied extensively in the last ten years (Ye et al., 2000; Giuliano et al., 2003; Lewinsohn et al., 2005). The production of rice, containing relatively high levels (2 mg/g) of provitamin A have been considered of great importance. Golden rice was created by transforming rice with two beta-carotene biosynthesis genes: *psy* (phytoene synthase) from daffodil (*Narcissus pseudonarcissus*) and *crt1* from the soil bacterium *Erwinia uredovora* (Ye et al., 2000). Due to various concerns about using genetically engineered crops, researcher now are trying the conventional breeding approaches for breeding varieties with increased beta-carotene in the aleurone.

Medicine

The promising applications of modern biotechnology in medicine cover such topics as pharmacogenomics; drug production; genetic testing; and gene therapy. *Pharmacogenomics* is the study of how variations in the human genome affect body's response to drugs. Such studies help to produce drugs, associated with specific genes and diseases.

Genetic testing includes the diagnosis of genetic disease and the detection of future disease risks.

Gene therapy

A "normal" gene is transferred into the genome to replace an "abnormal," disease-causing gene by using different types of viruses, such as retroviruses, adenoviruses, adeno-associated viruses and herpes simplex viruses. There are several non-viral techniques for gene delivery available now: direct introduction of therapeutic DNA into target cells; the use of liposome carrying the therapeutic DNA; chemical linking of the DNA to a molecule that will bind to special cell receptors.

Despite of developments in gene therapy research, there are many factors that create problems in effectively using gene therapy. Immune responses reduce possibility of gene therapy and the use of viral vectors causes toxicity and inflammatory responses. The main limitation of gene therapy is the fact that the most commonly occurring disorders are caused by combination of different genes and is difficult to treat using gene therapy.

Many reports and review articles demonstrate the latest developments in gene therapy for cancer and ther human diseases. It was reported that gene therapy is a promising approach for treatment of stroke and other cerebrovascular diseases (Chu et al., 2007).

Environment

The role of new biotechnology in solving environmental problems is very important. Bioremediation is one of the main branches of environmental biotechnology. Bioremediation can be characterised as any approach that apply microorganisms, fungi and green plants to restore the contaminated environment to its native condition. Various issues associated with water, air and soil pollution can be solved with new biotechnology. Nature evolved mechanisms for self-regeneration. The role of biotechnology is to efficiently apply these existing platforms to clean up environmental contamination. Different microorganisms are being examined for the ability to remediate various chemicals often present at polluted industrial sites. Researchers make the attempts to genetically modify certain microorganisms to increase their effectiveness to metabolize particular chemicals, such as hydrocarbons, in contaminated locations.

The future of biotechnology

Progress in biotechnology is aimed to achieve healthier planet, find solutions against deadly diseases. Stem cell research may find effective treatments for Parkinson's disease, multiple sclerosis, and muscular dystrophy. The future of biotechnology is directed to the improvement of the quality of life, finding solutions to cure different diseases and stop the hunger.

REFERENCES

Ashikari M., Sakakibara H., Lin S.Y., Yamamoto T., Takashi T., Nishimura A., Angeles E.R., Quian Q., Kitano H., Matsuoka M. (2005) Cytokinin oxidase regulates rice grain production. Science **309:** 741-745.

Botstein D., White R.L., Skolnick M., Davis

R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. **32:** 314-331.

Chu Y., Miller J.D., Heistad D.D. (2007) Gene therapy for stroke: 2006 overview. Curr. Hypertens Rep. **1:** 19-24.

Concibido V.C., La Valee B., McLaird P., Pineda N., Meyer J., Hummel L., Yang J., Wu K., Delannay X. (2003) Introgression of a quantitative trait locus for yield from *Glycine soja* into commercial varieties. Theor. Appl. Genet. **106:** 575-582.

Giuliano G., Al-Babili S., Von Linting J. (2003) Carotenoid oxygenases: cleave it or leave it. Trends Plant Sci. **8:** 145-149.

Gur A., Zamir D. (2004) Unused natural variation can lift yield barriers in plant breeding. PLoS Biol. **2:** e245.

Hacia J.G., Fan J.B., Ryder O., Jin L., Edgemon K., Ghandour G., Mayer R.A., Sun B., Hsie L., Robbins C.M., Brody L.C., Wang D., Lander E.S., Lipshutz R., Fodor S.P., Collins F.S. (1999) Determination of ancestral alleles for human single-nucleotide polymorphisms using high-density oligonucleotide arrays. Nat. Genet. **22:** 164-167.

Huang X.Q., Coster H., Ganal M.W., Roder M.S. (2003) Advanced backcross QTL analysis for the identification of quantitative trait loci alleles from wild relatives of wheat (*Triticum aestivum* L.). Theor. Appl. Genet. **106:** 1379-1389.

Kovacs M., Howes N.K., Clarke J.M., Leisle D. (1998) Quality characteristics of durum wheat lines deriving high protein from *Triticum dicoccoides* (6b) substitution. J. Cereal Sci. **27:** 47-51.

Lewinsohn E., Sitrit Y., Bar E., Azulay Y., Meir A., Zamir D., Tadmor Y. (2005) Carotenoid pigmentation affects the volatile composition of tomato and watermelon fruits, as revealed by comparative genetic analysis. J. Agric. Food Chem. **53:** 3142-3148.

Liu Z.J., Li P., Kocabas A., Ju Z., Karsi A., Cao D., Patterson A. (2001) Microsatellite-containing genes from the channel catfish brain: evidence of trinucleotide repeat expansion in the coding region of nucleotide excision repair gene RAD23B. Biochem. Biophys. Res. Commun. **289:** 317-324.

O'Brien S.J. (1991) Molecular genome mapping: lessons and prospects. Curr. Opin. Genet. Dev. **1:** 105-111.

Septiningsih E.M., Prasetiyono J., Lubis E., Tai T.H., Tjubaryat T., Moeljopawiro S., McCouch S.R. (2003) Identification of quantitative trait loci for yield and yield components in an advanced backcross derived from the *Oryza sativa* variety IR64 and the wild relative *O.rufipogon*. Theor. Appl. Genet. **95:** 418-423.

Welsh J., McClelland M. (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res. **18:** 7213–7218.

Williams J.G.K., Kubelik A.R., Livak K.J., Rafalski J.A., Tingey S.V. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. **18:** 6531–6535.

Vos P., Hogers R., Bleeker M., Reijans M., van de Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M., Zabeay M. (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. **23:** 4407-4414.

Xiao J.H., Grandillo S., Ahn S.N., McCouch S.R., Tanksley S.D., Liu J.M., Yuan L.P. (1996) Genes from wild rice improve yield. Nature **384:** 223-224.

Ye X., Al-Babli S., Kloti A., Zhang J., Lucca P., Beyer P., Potrykus I. (2000) Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. Science **287:** 303-305.