

## REVIEW ARTICLE

# Pharmacogenetics II. Research Molecular Methods, Bioinformatics and Ethical Concerns

E Daudén

Servicio de Dermatología, Hospital Universitario de la Princesa, Madrid, Spain

**Abstract.** Pharmacogenetics refers to the study of the individual pharmacological response based on the genotype. Its objective is to optimize treatment in an individual basis, thereby creating a more efficient and safe personalized therapy. In the second part of this review, the molecular methods of study in pharmacogenetics, including the microarray technology or DNA chips, are discussed. Among them we highlight the microarrays used to determine the gene expression that detect specific RNA sequences, and the microarrays employed to determine the genotype that detect specific DNA sequences, including the polymorphisms, particularly single nucleotide polymorphisms (SNPs). The relationship between pharmacogenetics, bioinformatics and ethical concerns is reviewed.

**Key words:** pharmacogenetics, DNA, molecular biology, microarrays, bioinformatics, genotype, gene expression, polymorphisms, SNP.

## FARMACOGENÉTICA II. MÉTODOS MOLECULARES DE ESTUDIO, BIOINFORMÁTICA Y ASPECTOS ÉTICOS

**Resumen** La farmacogenética es el estudio de la respuesta farmacológica del individuo según el genotipo. Su objetivo es optimizar el tratamiento a nivel individual, ir a una terapia personalizada más segura y eficiente. En esta segunda parte de la revisión se analizan los métodos moleculares de estudio en farmacogenética que incluyen la tecnología de los microarrays o chips de ADN. Dentro de ellos destacan los microarrays para determinar la expresión génica, que detectan secuencias específicas de ARN, y los microarrays para determinar el genotipo, que detectan secuencias específicas de ADN, incluyendo los polimorfismos, especialmente los de un solo nucleótido (SNP). Se revisa la relación entre la farmacogenética, la bioinformática y sus aspectos éticos. Palabras clave: farmacogenética, ADN, biología molecular, *microarrays*, bioinformática, genotipo, expresión génica, polimorfismos, SNP.

## Molecular Methods for Studying Pharmacogenetics: DNA Microarray Technology

Since the 1980s, the structure and expression of individual genes has been analyzed by techniques such as Southern blot analysis (hybridization with a DNA probe) and Northern blot analysis (hybridization with an RNA probe). These techniques were used to analyze a specific gene or small group of genes in each experiment. The problem is that to find associations between a gene and a

given disease, hundreds of thousands of permutations have to be tried in what would be a slow, laborious, and not to mention expensive process that often ended in failure. Study of expression of the complete genome gene by gene would be like trying to empty the ocean with a teaspoon. However, new techniques have been developed that allow thousands of genes to be analyzed in a single experiment.<sup>1</sup> In the future, these techniques will become routine. Currently, the most widely used technique for this type of study uses the so-called microarrays or DNA chips.<sup>2</sup>

The origins of microarray or DNA chip technology can be traced back to the early 1990s.<sup>3</sup> Since then, use of this technique in biomedical research has increased spectacularly, and expectations for the future are high. The technique is based on 2 complementary nucleic acid strands (specific target sequences of RNA or DNA are detected through specific binding of complementary sequences in a DNA probe).<sup>4</sup>

Correspondence:  
Esteban Daudén Tello.  
Servicio de Dermatología.  
Hospital Universitario de La Princesa.  
Diego de León, 62. 28006 Madrid. Spain  
E-mail: estebandauden@medynet.com

Manuscript accepted for publication October 23, 2006.

The microarray is a support that anchors gene fragments, oligonucleotides (short DNA fragments), or polymerase chain reaction (PCR) products. When this microarray is brought into contact with material from tissue samples taken from patients, it is possible to identify which genes are present through their complementarity. With this technique, it is possible to identify the genes present in normal or diseased tissues and compare the number of genes expressed and their level of expression in normal and pathological samples. Up to hundreds of thousands of DNA fragments with specific sequences (which together represent the genes of the genome several times over) can be placed on a few centimeters of special glass in well-defined, specific positions. These microarrays are manufactured using the same technology as used for semiconductor chips, but with millions of DNA strands deployed vertically on a glass or silicone chip. This is done using a technique known as combinatorial chemistry. Up to 1.3 million oligonucleotide probes can be synthesized on each array (each oligonucleotide is placed on a specific area of the array known as a probe cell, and each probe cell contains hundreds of billions of copies of a given oligonucleotide).

The design and manufacture of DNA microarrays depends initially on the target nucleic acid. This process requires the complete or partial characterization of the genome. Different genes can be identified through sequencing thousands of clones from libraries of genomic DNA and complementary DNA (cDNA). With this information, which is stored in large databases such as GenBank (<http://www.ncbi.nih.gov>) or the European Molecular Biology Laboratory (EMBL) (<http://www.ebi.ac.uk>), clones representative of each of the genes are identified. The DNA in each of these clones is amplified by PCR and purified for subsequent deposition, with precision robots, on the glass supports that comprise the DNA chips. Alternatively, small oligonucleotides (25-70 residues) are designed to hybridize specifically with gene or cDNA sequences. The oligonucleotides can be synthesized before immobilization on the supports or synthesized in situ on the chip. For this process, the different providers employ different methods: Affymetrix ([www.affymetrix.com](http://www.affymetrix.com)) uses photolithography; Agilent ([www.agilent.com](http://www.agilent.com)) injects phosphoramidite precursors; and Roche ([www.roche-applied-science.com/sis/matrixarray/](http://www.roche-applied-science.com/sis/matrixarray/)) activates precursors with an electric field.

Future developments will presumably offer cheaper chips specific for a given disease or drug response. These chips will be easier to use than the current devices, thereby opening the doors to widespread use in both the public and private health sectors. Although the sensitivity and specificity of these techniques remains the subject of debate,<sup>5</sup> they are currently beginning to come onto the market.

This technique allows the behavior of a gene to be analyzed, that is, the upregulation or downregulation of all genes expressed in a tissue as a result, for example, of a disease or treatment. Chromosomal losses or gains, polymorphisms, mutations, and changes in the levels of expression of the genes can be detected. With these findings, we can determine whether a person has a disease “pattern” or “profile,” is susceptible to developing a disease, or whether a person can respond to treatment.

Microarrays come in 3 different types:

1. Microarrays for determining gene expression: these detect specific RNA sequences.
2. Microarrays for genotyping: these detect specific DNA sequences.
3. Microarrays for resequencing.

Given that individual differences in response to treatments—the subject of this review—are determined by gene expression and genetic polymorphisms rather than the genes themselves, we will essentially focus on studies with the first 2 types of microarray.

## Analysis of Gene Expression

Only a fraction of the genes carried by an individual are expressed. Gene expression is the mechanism whereby the information contained in the DNA is transcribed into messenger RNA (mRNA), which is translated during protein synthesis. Gene expression allows the cell to adapt and respond to stimuli from other cells or the environment. Knowing which genes are the “active” ones is fundamental for determining the genetic profile in each situation. The pioneering study for determining gene expression using DNA microarray technology was done by Lockhart et al<sup>6</sup> in 1996. In that study, DNA oligonucleotides allowed a few mRNA molecules per cell to be identified. At the time, the DNA microarrays could detect the expression of around 1000 genes. Nowadays, microarrays include probes of all currently known genes (several thousand).

Gene expression is studied by measuring the number of RNA copies produced by a gene (active gene). Those genes which, although present, remain inactive are not detected. Analysis of gene expression with microarrays is often done blindly to start with. An extensive panel of genes is studied, and the results are then analyzed for possible influence on the disease or therapeutic response. With such an approach, the researcher starts without a specific hypothesis. It is possible to determine the expression of each of the known genes in the complete human genome, including those whose function is unknown. By simply comparing gene expression patterns (for example, patients with psoriasis who are responders or nonresponders to

cyclosporine), differences in these patterns can be detected. When a difference in these patterns of gene expression is detected, it makes sense to start a more targeted investigation.

Microarrays for determining gene expression are based on natural chemical attraction (known as hybridization) between DNA (on the array) and the RNA molecules (in the study sample) to determine which RNA sequences are expressed in a given sample and their level of expression from a certain gene (that is, which RNA and the amount of RNA produced). When a DNA strand produces an RNA strand, the two strands are complementary because their base pairs match up (A, T, G, and C in the DNA strand with U, A, C, and G, respectively, in the RNA strand). If the bases are not complementary, the DNA and RNA will not bind together (a single base that is not complementary is able to prevent 2 strands from binding). There are 2 types of microarray for determining gene expression. Currently, both types of microarray are similarly employed in research laboratories throughout the world:

1. DNA oligonucleotide microarrays. Small DNA fragments of chemically synthesized DNA (oligonucleotides) specific to each gene expressed are immobilized. Thanks to their homogeneity, reproducibility, and robustness, these are the most widely used. Several providers are available: Affymetrix (the main provider of this type of microarray; this company offers high density chips that, with more than 500 000 oligonucleotides, allow the simultaneous analysis of the expression of more than 20 000 genes) (Figure 1), Agilent, and Roche.
2. cDNA microarrays. These arrays use DNA probes of approximately 600 to 2000 bases, instead of 25,

anchored to a glass, nitrocellulose, or nylon support. cDNA fragments from collections of clones (libraries) are immobilized.

### Methodology in Oligonucleotide Microarray Techniques

The methodology used in oligonucleotide microarray gene expression techniques includes the following phases:

1. *Extraction and preparation of RNA from a sample (blood or tissue).* Total RNA is obtained from leukocytes in the blood samples and/or from tissue after grinding.
2. *RNA amplification and sequencing.* To facilitate hybridization, PCR amplification of the RNA (millions of copies) is done. The RNA strands are then broken down into millions of sections. Biotin molecules are also added to act as glue for fluorescent molecules.
3. *RNA identification.* Analysis of the expression of candidate genes is done with high-density oligonucleotide microarrays. This microarray examines practically every gene in the human genome. It contains more than 45 000 probe sets which correspond to 33 000 known genes and 6000 expressed sequence tags, that is, approximately the whole human genome. This allows identification of those genes, whether known or unknown, that are expressed, and also quantification of the level of expression of each of these genes.

In these microarrays, the DNA probes are placed on the surface of a glass support. Each probe is generally just 25 bases long—a small fragment but very representative of a complete gene with many more bases. If an RNA strand binds to these 25 bases, it is possible to accurately identify



**Figure 1.** DNA oligonucleotide microarray (Affymetrix, Santa Clara, USA).



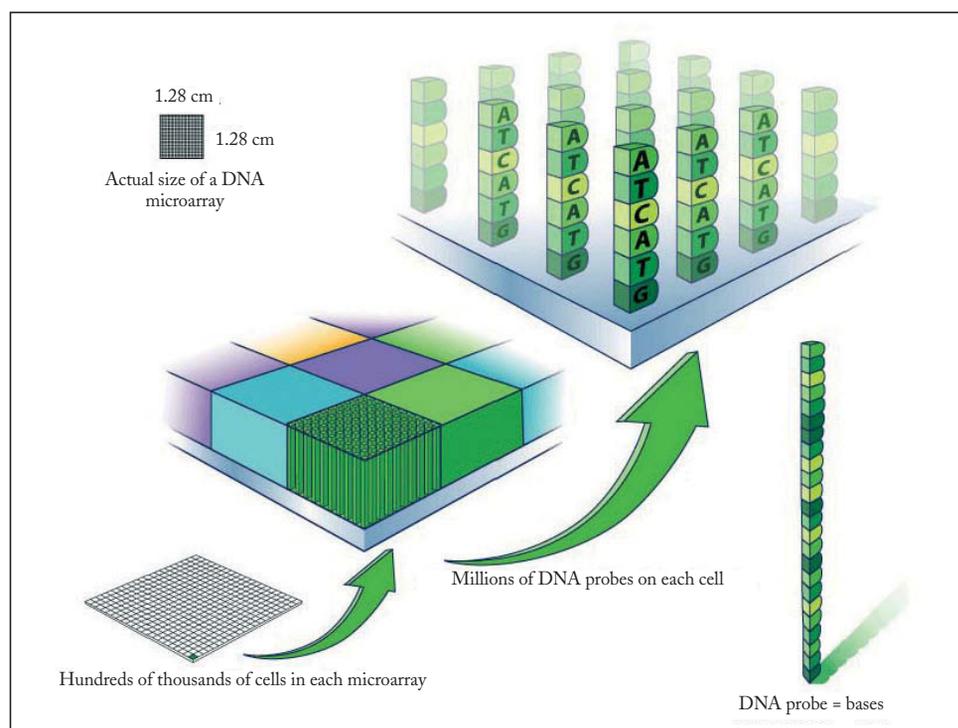
the gene. In fact, the microchip is so accurate that it can even detect a single specific RNA molecule of a gene from 100 000 different RNA molecules also present in the sample (sensitivity of detection of approximately 1:100 000).

The surface of this microarray is like a checkerboard measuring 1.25 by 1.25 cm. In total, the surface has up to 1.3 million cells. Each cell, or feature, measures 11 by 11  $\mu\text{m}$  and contains millions of copies of each probe of the same DNA molecule (Figure 2). The probes are stacked up next to each other by a photolithographic synthesis method.

The following steps are included in the process whereby RNA is identified: *a)* RNA placement on the microarray. The RNA extracted from the sample is amplified and deposited on the microchip. Between 14 and 16 hours are allowed to pass for hybridization to take place (so that RNA bases of the sample and the DNA of the probes on the microarray can bind to one another). Each RNA strand of genes expressed is floating among the DNA probes in search of its perfect complement. Unfortunately though, most RNA strands do not find their complement in the array. Those that do find their complementary sequence adhere to the DNA probe (Figure 3). *b)* Washing the microarray. The microarray is then washed to eliminate unbound RNA. RNA hybridized to the DNA probe remains, and is in turn bound to a biotin molecule. *c)* Observation of hybridized RNA. A fluorescent molecule that binds to biotin is used to visualize RNA bound to the DNA probe. After washing,

the unbound fluorescent molecules are eliminated to leave only those bound to biotin as markers of the RNA strands hybridized to DNA. Therefore, detection of fluorescent light will act as a marker of hybridization and allow detection of a specific sequence. The fluorescence is read with a scanning laser. The intensity of the fluorescence will give an indication of the relative quantity of each sequence of the sample: larger quantities of hybridized RNA will yield stronger fluorescence. If a gene is highly expressed and many RNA molecules are present, the fluorescence will be strong. Therefore, the amount of RNA can be quantified according to the intensity of the fluorescence. In turn, by scanning every cell, it is possible to evaluate all the several thousand genes present on the array. In short, knowing the identity of an oligonucleotide probe (and its corresponding gene) by its location on the microarray, we can identify the gene expressed, and from the intensity of fluorescence, determine the level of expression (Figure 4).

To compare different samples, a so-called heat map is constructed. After the laser scan, the features are displayed with different colors according to the level of expression. These maps are thus graphical representations in which gene expression is shown with colors, with strongly expressed genes being assigned intense colors. These maps can be recorded, printed out, or stored for evaluation at a later date. The image obtained in this way is analyzed by computer (Figure 5).



**Figure 2.** Composition of a DNA microarray.

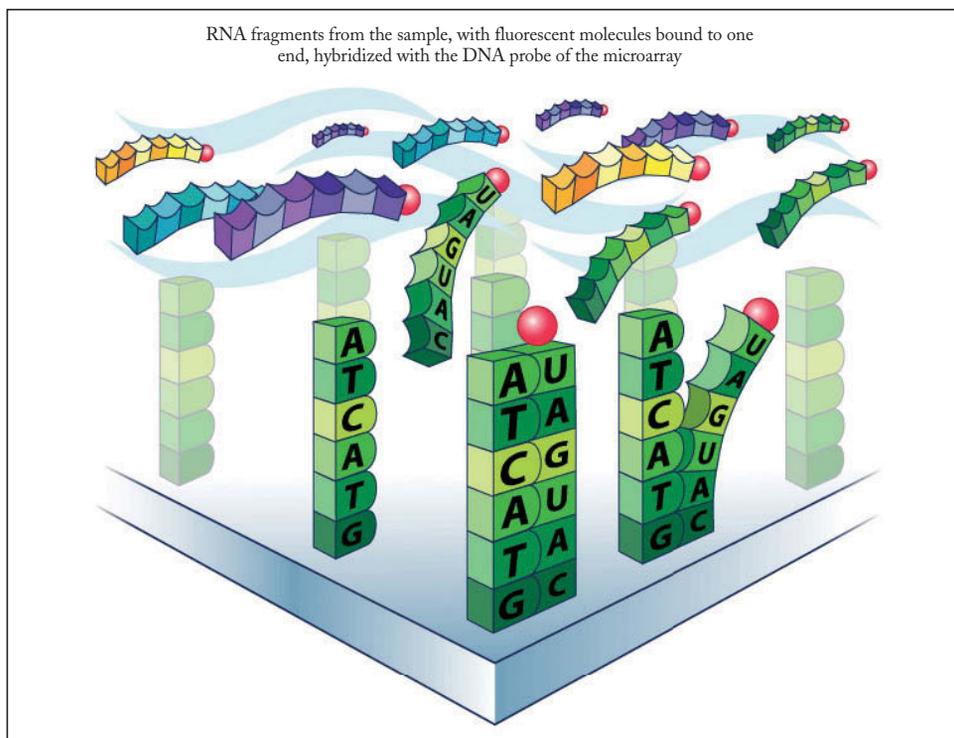


Figure 3. RNA hybridization.

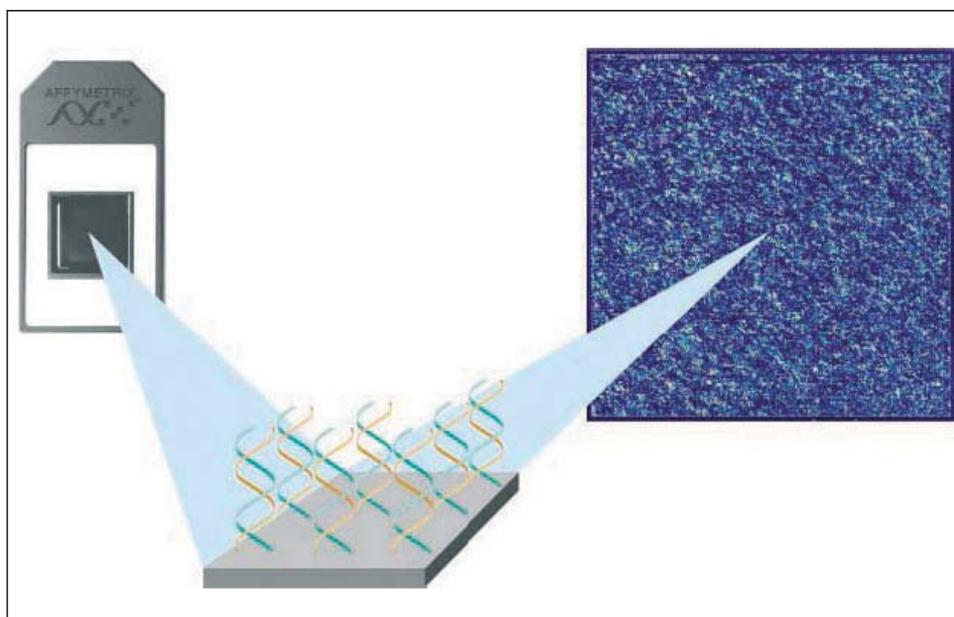


Figure 4. Result of RNA hybridization.

### Uses of Oligonucleotide Microarray Techniques

The uses of this technology are as follows:

1. *Disease classification.* If we compare the gene expression in different patients with a given disease, any differences we find could perhaps indicate disease subtypes. Patients with the same disease probably all
2. *Investigating pathophysiological mechanisms.* Once gene expression has been determined, a wealth of information can be obtained on the pathogenesis of a given disease provided we know the function of the gene.

express certain genes, but some might also express other disease-related genes, which would give a more detailed picture of the disease.

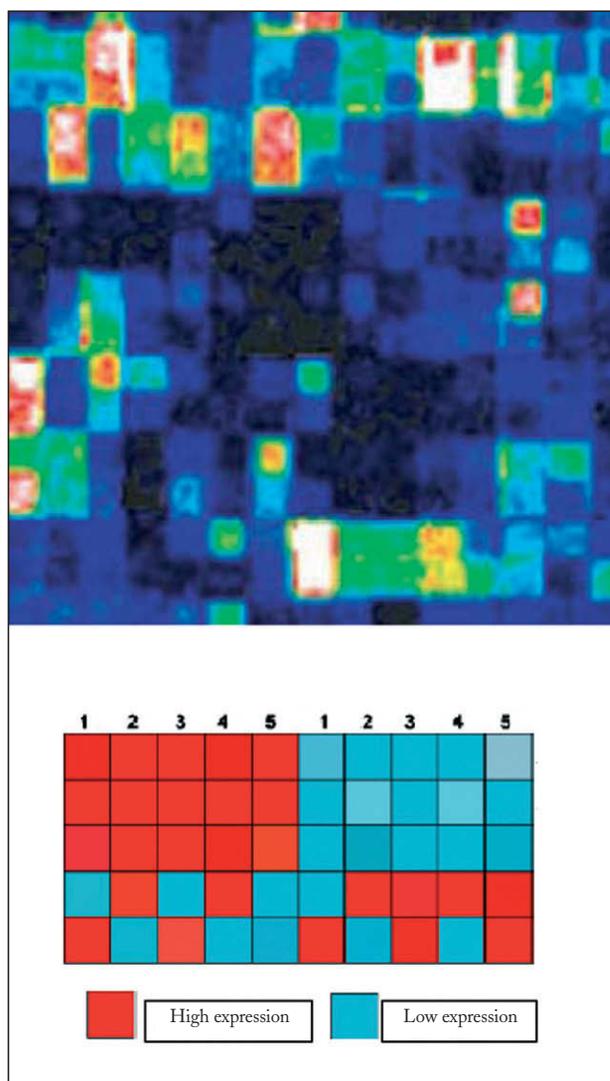


Figure 5. Color map.

3. *Drug discovery.* In accordance with the preceding section, knowledge of the pathogenesis of a disease can help the search for new drugs that target the mechanisms in which a particular gene participates. The new technology is also useful, however, not just for giving an indication of efficacy but also for determining toxicity.
4. *Predicting the therapeutic response of patients to drugs and optimization of their use: tailored therapy.* As discussed previously, patients with a certain common disease may respond differently to the same drugs. Comparison of gene expression in a patient before and after receiving a drug should reveal the differences present. If patients who respond have a characteristic pattern that differentiates them from nonresponders, we can predict which patients will or will not benefit from a drug. In

this way we can assess efficacy. By carrying out this process with different doses of a drug, we can also determine what doses are necessary for certain patients (for example, those with a certain gene expression pattern who fail to respond at normal doses but become responders at higher doses). Such an approach can help tailor and adjust the necessary dose. The same can be said of toxicity; we are able to predict which patients will suffer side effects and which will not.

## Genotyping

Although the genome sequence is mostly homogenous, around 1 in every 100 to 1500 nucleotides are polymorphic, that is, they have a base on one chromosome that is different to the corresponding base on the other chromosome. These differences, known as polymorphisms, have consequences for protein expression or structure. Thus, in practice, only monozygotic twins have a highly similar genetic profile, whereas there are notable differences between individuals of the human race. Each person inherits 2 copies of each gene, 1 from each parent. These copies can be identical or different. Taken together, these inherited differences are responsible for the genetic variation between individuals. Very small differences in DNA sequence can have large effects on health and diseases. A gene that functions correctly in one person may only have a reduced function or even not function at all in another. Genetic polymorphisms can be defined only when single-gene variants are present in more than 1% of the normal population.<sup>7</sup> The types of polymorphism are as follows:

1. Serious mutations in DNA, with noteworthy repercussions. These might include *a)* deletions (elimination of 1 or more bases or segments from the DNA sequence). Deletions usually have a substantial negative impact on the implicated gene and the protein it encodes. *b)* Insertions (an extra base or sequence of bases is added to the DNA sequence). This may or may not take place in the coding region of the gene. Insertions generally have an extremely negative effect on the encoded proteins.
2. Minor mutations, with less serious repercussions. These include single nucleotide polymorphisms (SNPs), the most common type of polymorphism.

SNPs are homo- or heterozygotic variants of a single base pair in the DNA sequence of the human genome. These are the simplest and the most common form of polymorphism between individuals, and are one of the most powerful tools for analyzing genomes. Most SNPs are located outside the coding regions or promoter of the

gene, but regardless of whether they are inside or outside these regions, SNPs can still influence gene transcription. The most important characteristics of SNPs—and those which make them ideal genetic markers in the search for genes indicative of susceptibility to a disease or response to a given drug—are their simplicity, widespread distribution, stability, and high frequency of occurrence throughout the genome (it is estimated that an SNP is present in every 1000 base pairs).<sup>8</sup>

More than 1.4 million SNPs have been identified in the initial sequencing of the human genome (with more than 60 000 of them in coding regions),<sup>9</sup> and it is estimated that the human genome contains around 10 million SNPs. The SNP Consortium (a nonprofit organization made up of pharmaceutical companies, information processing companies, academic institutions, and research foundations) has produced a high density map with several million SNPs (<http://snp.cshl.org>). Thanks to this map, thousands of polymorphisms can be accurately identified and classified.

The following steps were used in the past and continue to be followed to achieve the goals of pharmacogenetics: *a)* analysis of the distribution of the target variant in the population using a probe to detect a phenotypic polymorphism (affecting >1% of the population); *b)* identification of the gene responsible and its variants; *c)* familial and twin studies to confirm the genetic nature of the variant (dominant or recessive, Mendelian inheritance, etc); *d)* development of genetic assays for the different DNA variants; *e)* correlation between genotype and phenotype; and *f)* application in clinical practice.

In fact, most pharmacogenetic studies published to date are based on correlating the response of patients with variations in genes implicated in the mechanism of action of the drug or its metabolism in order to detect the “candidate gene.” On the other hand, over the last 10 years, advances in molecular sequencing technology have also made possible systematic screening to detect significant functional variants in DNA sequences of genes that influence the effects of a number of drugs for subsequent study of the phenotypic implications. Today, this is possible thanks once again to the use of DNA microchips for detecting SNPs.<sup>8-10</sup> Foreknowledge of the gene responsible is not essential given that it can be determined by analyzing differences between responders and nonresponders. This approach is based on complementarity, on the attraction of one DNA strand for another (and on the combination of base pairs: A, T, C, and G with T, A, G, and C, respectively). For genotyping, almost all chips are based on oligonucleotides. As before, Affymetrix is the main provider, with chips that allow thousands of known SNPs to be analyzed in a single experiment (at present, microarrays are being developed

with up to 100 000 SNPs). The Human Genome Project has managed to sequence a large portion of the human genome, and the results of this project are reflected in the microarray. The microarray contains DNA sequences (short but representative sequences of the genes of the genome) that allow detection of the full genotype (including SNPs) of an individual who provides a sample.

Determination of the genotype of an individual only needs to be done once for a given gene because, except for certain rare mutations, it does not usually change (unlike analysis of RNA expression, in which gene expression and hence RNA production varies according to cell type). Genotyping can be done with any type of tissue or fluid, because the DNA is the same for all cells in the organism.

### Microarray Techniques for Detecting Single Nucleotide Polymorphisms

The microarray techniques used for detecting SNPs are similar to those described for gene expression:

1. *Extraction and amplification of DNA from a sample (blood or tissue).*
2. *DNA amplification and sequencing.* In order to facilitate detection of SNPs, an amplification (millions of copies) of each DNA fragment containing these polymorphisms in the sample is undertaken by PCR. The DNA is then sequenced.
3. *Identification of the genotype.* The previously obtained DNA sample is placed on the microarray for detection of SNPs. The structure of the microarray is similar to that used in the RNA analysis; DNA oligonucleotide probes with 25 bases are stacked in hundreds of thousands of cells or features (each feature with a single type of DNA probe). This microarray includes most of the genes in the human genome and is able to analyze thousands of known SNPs in a single experiment. The process of DNA identification follows the same phases as for RNA:

- a)* The RNA is placed on the microarray.
- b)* The microarray is washed (to leave hybridized DNA).
- c)* Hybridized DNA is detected (Figure 6).

Fluorescent molecules are also used that bind to biotin, which in turn is bound to the DNA probes. The intensity of fluorescence after a laser scan is measured to build up the color maps. This process identifies the genes and polymorphisms present in the patient sample. By scanning all the cells of the array, it is possible to evaluate every one of the several thousand genes present on the array. We can identify the genes and the polymorphisms present because the identity of the oligonucleotide probes (and of course

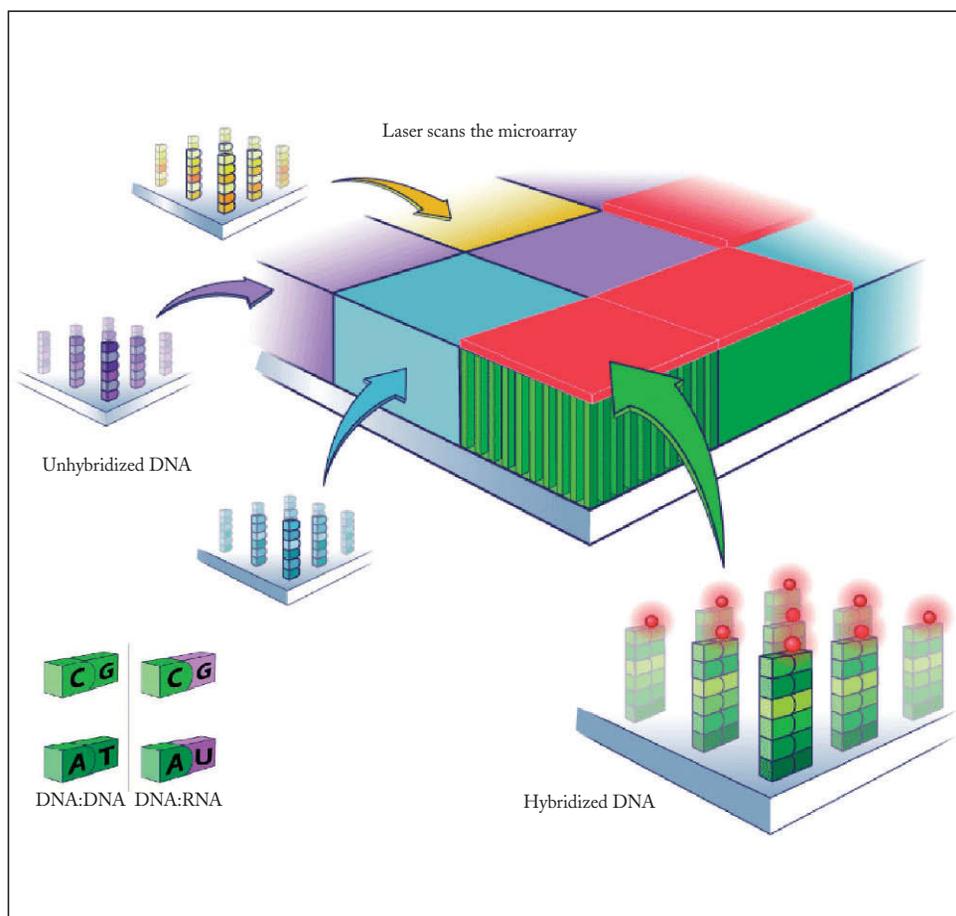


Figure 6. DNA hybridization.

their corresponding genes) is defined by their location on the microarray. It is also possible to determine whether an SNP is homozygous or heterozygous according to the hybridization pattern.

### Aims of Polymorphism Studies

Polymorphisms are studied for the following reasons:

1. *To analyze genetic markers of disease susceptibility.* Analysis of polymorphisms can indicate whether an individual is predisposed to a certain disease. The candidate gene or group of genes that contribute to the development of the disease can be identified.
2. *To determine the predisposition of an individual to a certain subtype within a disease.* For diseases with a wide range of clinical presentations, such as psoriasis, where the genetic profile may be partly responsible for the observed clinical variability.
3. *To study the development and progression of tumors.* SNPs have been implicated in the biology of tumors.
4. *To identify therapeutic targets.* As in the analysis of gene expression, genotyping can also contribute to pharmacogenomics.
5. *To predict therapeutic response to a drug.* The variability in genetically determined pharmacological response is closely linked to polymorphisms in specific genes. According to a number of studies, some SNPs have been associated with changes in metabolism, transporter proteins and drug receptors, degree of response to certain drugs, and level of toxicity. In fact, some SNPs are already being used to predict clinical response to drugs.<sup>11-14</sup> Individuals who carry a certain allele are also probably carriers of specific variants with several markers of SNP. Therefore, to predict pharmacological response, it may not be necessary to identify the actual genes and alleles implicated as detection of SNP could be sufficient.

Future advances in robot technology and use of multiple reactions will significantly lower the cost of analysis of several hundred thousand SNP per patient in clinical trials.

With bioinformatics, these SNP patterns can be rapidly compared in patients with phenotypic differences in drug responses. Thus, when information on the differences in pharmacological response is clearly defined, the SNP patterns can be compared with the diagnostic chips, allowing determination or prediction of the response of a given patient to a drug.

## Gene Resequencing

It is possible to identify DNA in a sample from an organism or even from a virus by gene resequencing. This is a tool for identification (for example, to determine the strain of a virus in an epidemic, locate pathogens in water or food samples, and determine whether a food label is accurate or whether the food has been adulterated). Such analyses are fairly complicated and are beyond the scope of this review.

## Pharmacogenetics and Bioinformatics

Bioinformatics is the “organization, manipulation, and analysis of results from molecular biology and biomedical experiments using computer-based methods,” that is, imposing cognitive order on the information accumulated. Although bioinformatics was born 30 years ago, its importance and central role in biomedicine is clearly linked to the proliferation of genomic techniques. The development of bioinformatics has gone hand in hand with the needs generated by the rapid progress in experimental techniques. This is particularly true in the case of molecular genetics, which from the outset has required experimentally generated information on gene sequences to be organized, stored, and analyzed.<sup>15</sup>

Essentially, bioinformatics provides a way of tackling the problem of analyzing the sequences of the nucleic acids contained within genomes. Currently, databases house 31 million entries, 45 billion bases, and a million protein entries with 310 million amino acids. Such a volume of data poses computational challenges. In addition, the volume of data produced by molecular biology and biomedicine is not only large but also very diverse. This situation poses difficult problems for designing the corresponding database structures (storage methods), structures which have to be continually adapted to the new methods.

The 2 main centers that house essential databases are the National Center for Biotechnology information (NCBI) in the United States of America and the European Bioinformatics Institute affiliated to EMBL (EBI-EMBL) in Europe, although many other databases are available that concentrate on a range of

aspects. The main databases in molecular biology and biomedicine include:

1. National Center for Biotechnology information (NCBI). This database is part of the National Institute of Health National Library of Medicine ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)).
2. European Bioinformatics Institute (EBI-EMBL) ([www.ebi.ac.uk](http://www.ebi.ac.uk)).
3. Sanger Center ([www.sanger.ac.uk](http://www.sanger.ac.uk)).
4. Protein Design Group (PDG, CNB-CSIC) ([www.pdg.cnb.uam.es](http://www.pdg.cnb.uam.es)).
5. DNA sequences ([www.ebi.ac.uk/embl](http://www.ebi.ac.uk/embl)).
6. Protein sequences ([www.ebi.ac.uk/embl](http://www.ebi.ac.uk/embl)).
7. Protein domains ([www.ebi.ac.uk/interpro/index.html](http://www.ebi.ac.uk/interpro/index.html)).
8. Three-dimensional structures ([www.ebi.ac.uk/MSD](http://www.ebi.ac.uk/MSD)).
9. Molecular basis of diseases ([www.ncbi.nlm.nih.gov/omim](http://www.ncbi.nlm.nih.gov/omim)).
10. A system for navigating between databases (<http://srs.ebi.ac.uk>).

On the other hand, several public and private organizations provide methods for analyzing the results of DNA microarrays, including the creation of standards for comparing the results generated by different laboratories. In the specific case of pharmacogenetics, bioinformatics can provide *a)* organization and storage of the experimental information generated; *b)* design of databases adapted to the technology used in pharmacogenetics; *c)* capacity for storing both images and data obtained from the studies performed; *d)* estimation of the statistical significance of the results; *e)* comparison of expression patterns of different samples (from the same patient or between patients); and *f)* analysis of results including comparison with information on function, diseases, and drugs available, to predict the efficacy of a medication and its tolerance in a given patient with a given statistical reliability.

## Pharmacogenetics and Ethical Considerations

The use of genetic tests to diagnose or predict illness, to predict the response to treatment, or to determine possible adverse effects of treatment will soon become available in clinical practice as a tool for improving the health of individuals on the basis of their genetic susceptibility to certain diseases or predicted therapeutic response. However, the information that might be obtained as a result of performing these genetic tests raises questions about the access to that information and the ends for which it is used because conflicts of interest may arise between the tested individual and other individuals or

entities. The uneasiness centers on deciding by whom, under what circumstances, and with what purpose genetic testing can be done after biological samples for the tests have been obtained. Once the tests have been done, there is further controversy concerning who will have access to the results, who can be informed of the results and under what circumstances, what uses can be given to the results, and what measures should be taken to protect the privacy of those tested.<sup>16</sup> These questions generate an ethical debate the scope of which, although centered on diagnostics and prediction of risk of disease (with implications in the health insurance sector, work, etc), also covers pharmacogenomics and pharmacogenetics.

A legal framework is therefore needed that addresses these questions. The Spanish state has already taken steps in this direction by ratifying the Convention for the Protection of Human Rights and Fundamental Freedoms in relation to the applications of biology and medicine. This ratification, known as the Oviedo agreement because it took place in Oviedo, Spain, on April 4, 1997 (Official State Journal [BOE] dated October 20, 1999), clearly establishes that the interests and well-being of human beings should take precedence over the interests of society and science. Given the need to ensure that individual interests and fundamental freedoms take precedence, a genetic test must conform to the requirements of the general health legislation passed in 1986. Among other things, this legislation requires informed consent, which should include information on the nature of the analysis, the aims thereof, and the consequences and procedure to adopt in the event of a positive result. Currently, informed consent establishes that the data obtained are controlled by the signatory. Individuals have the right to decide for themselves how the data are used and who can access them, to control whether they should be deleted and ensure their accuracy, and to authorize whether they are made available to third parties. This firmly establishes the precedence of the right of genetic privacy over any other right. No individual or public or private entities should have the right to analyze the genetic information of an individual without consent. The owner of genetic information is the individual. Therefore, any decision relating to transmission of that information to third parties, including genetic relatives, should be a strictly personal matter.<sup>17</sup>

Spain's Organic Law 15/1999 on the protection of personal data extended protection to all types of data, and not just data on electronic support. Protection therefore now covers any material of biological origin that contains identifiable personal information, thereby regulating computer-based genetic information. This means that if genetic records are established, storage of data should be voluntary and the confidentiality of these data should be guaranteed. System administrators should guarantee that it

is impossible to access, destroy, or disseminate genetic and personal data without the express consent of the individual to whom they belong.<sup>18</sup>

Furthermore, both the European Convention (April 4, 1997) on human rights and biomedicine (articles 5, 10.1, and 12) and the Universal Declaration of the Human Genome and Human Rights (November 11, 1997) of the United Nations Educational, Scientific and Cultural Organization (UNESCO) (article 5b) establish that prior, free, and informed consent will be obtained from all individuals as a part of their fundamental rights.<sup>19</sup> In this context, the UNESCO published an international declaration in 2003 on the protection of human genetic data, in order to avoid discriminatory use of information on the genes associated with predisposition to a range of diseases. This declaration could be extended to include therapeutic response. According to these recommendations, in recent years, most European countries have been implementing a series of guidelines and recommendations for the correct use of genetic tests.

Finally, the medical, ethical, and legal aspects concerning the use of genetic tests should be reviewed and updated periodically to keep in step with scientific progress and the changing face of society. The scientific community should play a key role in raising awareness in patients, and therefore society, regarding the correct use of these new technologies.

#### Conflict of Interest

The author declares no conflicts of interest.

## References

1. Shoemaker DD, Schadt EE, Armour CD. Experimental annotation of the human genome using microarray technology. *Nature*. 2001;409:922-7.
2. Kunz M, Ibrahim SM, Koczan D, Scheid S, Thiesen HJ, Gross G. DNA microarray technology and its applications in dermatology. *Exp Dermatol*. 2004;13:593-606.
3. Fodor SP, Read JL, Pirrung MC, Stryer L, Lu AT, Solas D. Light-directed, spatially addressable parallel chemical synthesis. *Science*. 1991;251:767-73.
4. Southern E, Mir K, Shchepinov M. Molecular interactions on microarrays. *Nat Genet*. 1999;21:5-9.
5. Lander ES. Array of hope. *Nat Genet*. 1999;21 Suppl: 3-4.
6. Lockhart DJ, Dong H, Byrne MC. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol*. 1996;14:1675-80.
7. Meyer UA. Genotype or phenotype: the definition of a pharmacogenetic polymorphism. *Pharmacogenetics*. 1991;1:66-7.
8. Shastri BS. SNP alleles in human disease and evolution. *J Hum Genet*. 2002;47:561-6.
9. Sachinadandam R, Weissman D, Schmidt SC. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*. 2001;409:928-33.

10. Hacia JG, Fan JB, Ryder O. Determination of ancestral alleles for human single-nucleotide polymorphisms using high-density oligonucleotide arrays. *Nat Genet.* 1999;22:164-7.
11. McLeod HL, Evans WE. Pharmacogenomics: unlocking the human genome for better drug therapy. *Annu Rev Pharmacol Toxicol.* 2001;41:101-21.
12. Roses AD. Pharmacogenetics. *Hum Mol Genet.* 2001;10:2261-7.
13. Evans WE, Relling MW. Pharmacogenomics: translating functional genomics into rational therapeutics. *Science.* 1999;286:487-91.
14. Evans WE, Johnson JA. Pharmacogenomics: the inherited basis for interindividual differences in drug response. *Annu Rev Genomics Hum Genet.* 2001;2:9-39.
15. Valencia A. Bioinformatics: Biology by other jeans. *Bioinformatics.* 2002;18:1551-2.
16. Ruiz J. Test genéticos: implicaciones éticas y jurídicas. *Spainfo SA;* 2004. p129-46.
17. Romeo Casanova CM. El tratamiento y la protección de los datos genéticos. In: Mayor Zaragoza F, Alonso Bedate C, editors. *Genética.* Madrid: Ariel; 2003.
18. Moreno RF. Información genética (informgenics) e intimidad. *El Derecho ante el Proyecto Genoma Humano.* Madrid: Fundación BBV; 1994.
19. Guttmacher A, Collins F. Ethical, legal and social implications of genomic medicine. *N Engl J Med.* 2003;349:562-9.