Array-based comparative genomic hybridization as a tool for solving practical biological and medical questions

David Blesa, Sandra Rodríguez-Perales, Sara Alvarez, Cristina Largo, and Juan C. Cigudosa

3.1 Introduction

Genetic aberrations, as losses of genetic material (deletions) or localized gains that affect certain regions of the genome, have been shown to be the basis of many diseases or human pathologies. Rare diseases, such as developmental abnormalities or mental retardation, or much more prevalent pathologies, such as cancer, are characterized by the occurrence of one or more of such genetic alterations in the genome that lead to changes in DNA sequence copy number.

Comparative genomic hybridization (CGH) has been one of the methods for the identification and further characterization of these genomic copy number changes. CGH is a molecular cytogenetic technique that allows the analysis of DNA gains and losses in the entire genome in a single hybridization experiment. It is based on the co-hybridization of two differentially fluorescence labeled DNAs to normal human metaphase chromosomes. Equal amounts of the labeled test and reference DNAs compete to hybridize proportionally to the copy numbers of the sequences present in the target chromosomes. In this way, the relative fluorescence intensity of the test to reference is determined along the length of the target chromosomes and differences between the abundance of complementary sequences in the hybridized DNAs are localized and quantified. This technique was developed by Kallioniemi *et al*. (1992) and, since then, it has contributed to the knowledge of the chromosomal aberrations present in many constitutional diseases and tumors. The sensitivity of the CGH technique depends on the degree of condensation of the chromosomes and on the size of the chromosomal aberration, something that limits

CGH's power of resolution to approximately five to ten megabasepairs (Mb) of DNA sequence. Changes (deletions, gains or amplifications) that affect genomic regions smaller than this size are not readable or efficiently detected by chromosome CGH.

Through the introduction, use and management of genome-based tools, research into genetic alterations that give rise to these diseases, as common as cancer, has undergone a technical revolution comparable to the advance of microscopy in the laboratory. Now, the study of the gene–disease relationship can be achieved by analyzing the behavior of thousands of genes, the complete genome if possible, in a simultaneous form. These systems, generically called arrays are changing the way we pose problems and draw conclusions from experiments, since they offer us a complex picture of the genome as a whole.

This change to genome-based approaches has had an immediate effect also in chromosome CGH. Metaphase spreads are being replaced as targets for hybridization by genomic microarrays. The limitations in the power of resolution of the chromosome CGH have been easily overcome by substituting the chromosome by small fragments of DNA arrayed onto a solid support: large-insert clones (BAC/PAC clones, i.e. bacterial artificial chromosomes, ~150 kb in length), complementary DNA (cDNA) clones, or oligonucleotides (Solinas-Toldo *et al.*, 1997; Pinkel *et al.*, 1998; Pollack *et al.*, 1999; Lindblad-Toh *et al.*, 2000; Lucito *et al.*, 2000; Mei *et al.*, 2000). Resolution is now limited by the type, amount, and distribution through the genome of the clones that are included in the array. Typically, most studies utilize whole genome microarrays comprising large-insert BAC or PAC clones spaced at approximately one clone per megabasepair (Fiegler *et al.*, 2003a), but higher resolution arrays comprising overlapping clone sets from specific regions (Buckley *et al.*, 2002) are also being employed. There are published experiments with array platforms that include from a few hundred clones to over 30 000 clones covering the complete human genome (Carter and Vetrie, 2004; Ishkanian *et al.*, 2004). The jump from chromosomes to clones as targets for CGH has also changed the name of the technique that can now be frequently mentioned as array based CGH, matrix-CGH, or, simply, array CGH.

The nature of the genetic aberrations that take place in diseases such as cancer and genetically determined mental retardation, that is amplifications, genomic gains and/or deletions, makes array CGH the most adequate approach to investigate them. An important advantage is that array CGH requires only DNA, which can be isolated from routine paraffin embedded pathology samples. This allows for studies on samples that have been stored for many years.

In this chapter, we present data about the use of different array CGH platforms to unveil the genomic abnormalities that can take place in two pathological conditions. First, we describe the use of array CGH to search for possible DNA copy number changes involving the subtelomeric regions of acute myeloid leukemia (AML) cells with apparently normal karyotype. Second, we describe the use of array CGH to speed up the process of cloning a familial chromosome translocation, t(3;8)(p14;q24), which is associated with the onset of renal cancer in those members of the family that are carriers of the chromosome aberration.

3.2 Scientific background

3.2.1 Leukemic cells with normal karyotype may show cytogenetically undetectable DNA copy number changes

The first case study deals with AML and normal karyotype. AML is a type of hematological tumor characterized by the proliferation of undifferentiated myeloid precursor stem cells (Huntly and Gilliland, 2005). The proliferating clone replaces normal stem cells production in the bone marrow resulting in a defective hematopoietic homeostasis with severe clinical consequences. This proliferation and blocked differentiation is commonly sustained or caused by a genetic molecular or chromosomal mutation that can be detected for proper diagnosis and monitored for therapeutic purposes. In fact, there is much information regarding the chromosomal changes that take place in AML: specific chromosome translocations, deletions, trisomies, for example (Heim and Mitelman, 1995). Most of these genetic aberrations are well characterized, even at the gene rearrangement level. They can be detected in a routine cytogenetic analysis (the karyotype) or, when the involved gene is known, by a fluorescence *in situ* hybridization (FISH) assay. These chromosome rearrangements may be used as prognostic factors and, in fact, they are currently used in clinical practice to classify patients in risk groups. Successful therapies for AML greatly rely on the correct classification of a patient in a determined group risk (Lowenberg, 2001). However, it is accepted that around 50% of the *de novo* cases of AML do not show chromosome rearrangements that are reliably detected by conventional cytogenetics or FISH assays. With the large amount of information that came from the Human Genome project and the availability of genomic analysis systems, we should explore the presence of other genetic aberrations that may take place in AML and which are not disclosed by the cytogenetic analysis due to its low resolution power. A CGH array works with genomic DNA and is then a very adequate tool to investigate gains (amplifications) and losses (deletions) that may be present in the target samples.

3.2.2 The cloning of a familial translocation associated with renal cell carcinoma

Renal cell carcinoma (RCC) comprises a heterogeneous group of tumors that have been divided into different subtypes based on histological features. Clear-cell RCC (CC-RCC, also known as nonpapillary RCC) is the most common type (75% of all RCCs). Although CC-RCCs mostly occur in a sporadic form, several familial cases have been reported. The most common form of familial CC-RCC is in association with the dominantly inherited von Hippel–Lindau (VHL) cancer syndrome. The other form is composed of families that show segregation of CC-RCC with constitutive balanced translocations involving chromosome #3. Previously, we described a Spanish family carrying a constitutional t(3;8)(p14.1;q24.32) translocation (Melendez *et al.*, 2003) (*Figure 3.1*). All CC-RCC patients in this family were carriers of this rearrangement. We speculated that deregulation of a gene(s) located at or near the translocation breakpoints may play

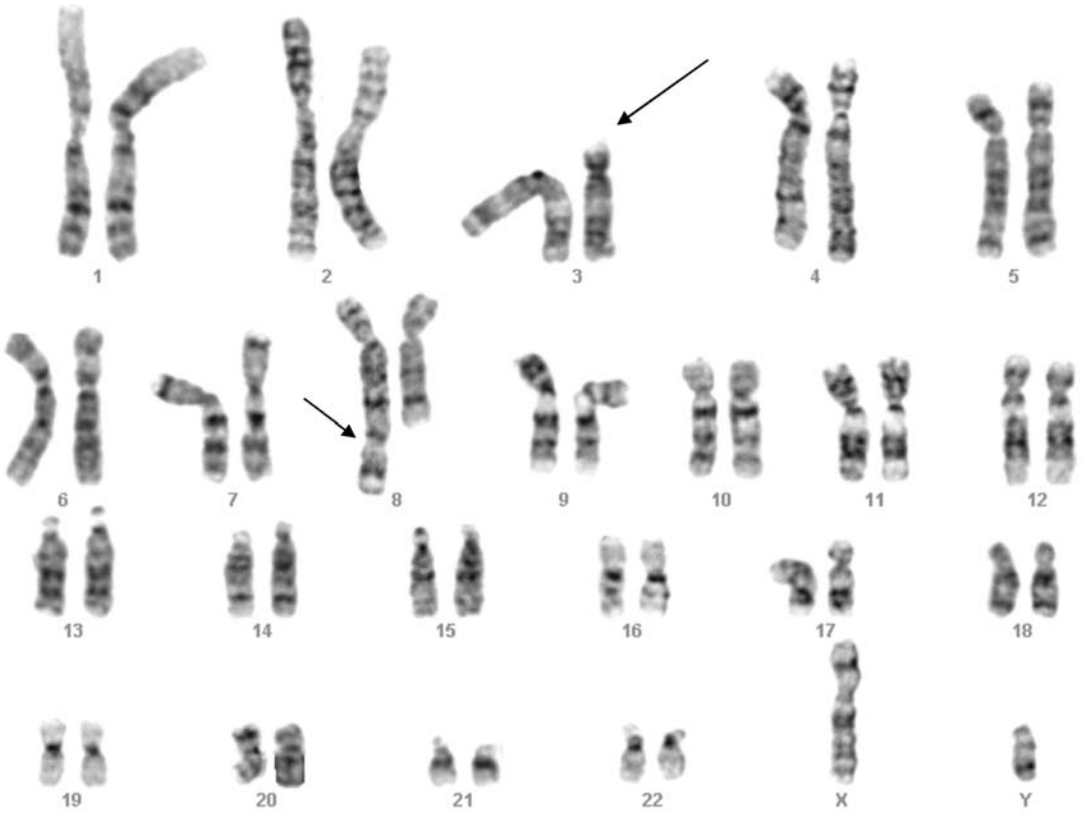

Figure 3.1

G-banded karyotype of a member of the family carrying the t(3;8)(p14.1;q24.32) translocation. Arrows identify the derivative chromosomes #3 and #8.

a role in the development of RCC in this family. A scientific project was then conducted to clone the chromosomal breakpoints that were involved in the translocation and to identify or detect the existence of any genes that may have been affected as a result of this rearrangement. This hypothesis was supported by the existence of breakpoint spanning genes with biological significance that are disrupted in some previously reported familial translocations involving chromosome #3 and associated with RCC: *FHIT* (located at 3p14), *TRC8* (8q24.1), *DIRC1* (2q33), *DIRC2* (3q21), *DIRC3* (2q35), *LSAMP* (3q13.3), and *NORE* (1q32.1) (Rodriguez-Perales *et al.,* 2004).

3.3 Design of the experiments

The alternative possibilities to study the presence of genomic DNA gains and/or losses in a given sample come from the different available platforms, and their specific features, that may be used in array CGH experiments.

The panel of platforms can be initially divided into three groups regarding the type and size of the clones arrayed: large BAC/PAC clones of a mean

size of 150 kb of genomic sequence, cDNA clones with sizes around hundreds of basepairs (bp) covering the coding sequence of genes, and oligonucleotides with sizes around 40–60 bp. Each type of platform has its own advantages and limitations. In recent years we have observed in the literature and in scientific forums a competition among providers and users of different platforms in order to demonstrate which one was the most powerful or efficient, or both, to detect DNA gains and losses. Regardless of economic or other reasons, there is a substantial amount of published work on this issue that allow us to draw some simple recommendations.

BAC arrays, also called genomic arrays, were the first introduced platform (Solinas-Toldo *et al.*, 1997). They are the most commonly used for their robustness and good signal in the hybridization experiments. Because of this good hybridization yield, they are especially useful in the detection of genomic losses (deletions) and of alterations affecting a single element of the array. Their two main disadvantages are that the DNA production for spotting onto the array is expensive and that each element (i.e. a BAC or PAC clone) usually contains several genes. If the research involves the localization of altered genes, each gene within the altered BAC/PAC clone has to be tested (Albertson and Pinkel, 2003; Fiegler *et al.*, 2003a, 2003b; Carter and Vetrie, 2004).

cDNA arrays used as CGH arrays are a second option. In this platform the clones are expressed sequences that have been obtained from a previously defined library. cDNA arrays, originally developed and extensively used for expression profiling, are not a first choice for studying copy number changes. They are widely available but were not intended for CGH analysis and yield poor signal to noise ratios for many clones and require a large amount of DNA (~10 μg) for hybridization. They are indicated for certain experiments for localization of genes that are simultaneously overexpressed and amplified. However, the cDNA approach is clearly not recommended if the objective is the detection of small altered regions or single copy gains or losses (Monni *et al.*, 2001; Hyman *et al.*, 2002; Pollack *et al.*, 2002; Clark *et al.*, 2003; Hedenfalk *et al.*, 2003).

Finally, oligonucleotide-based CGH arrays are the most recent approach. It implies the use of short sequences of new synthesized fragments of DNA (oligonucleotides) of 40 or 60 bp in length (40-mer or 60-mer oligos) as targets for hybridization in the slides. The main advantages of this type of array are the high coverage density that can be easily achieved (it is rather normal that they contain 40 000 clones), each oligonucleotide can be designed to yield the best possible hybridization result, can cover poor gene regions, and all genes can be represented in the array at one time. The main disadvantages are the high cost (which eventually will decrease) and, more importantly, the variable and not so robust yield of the hybridization signal that will require some improvements of the protocols and analysis software. They also require large amounts of DNA for hybridization and are not reliable for single element alterations (although their high density design overcomes, in part, this problem) (Barrett *et al.*, 2004; Bignell *et al.*, 2004; Huang *et al.*, 2004; Rauch *et al.*, 2004; Wong *et al.*, 2004; Herr *et al.*, 2005).

The old controversy about the low density of the BAC genomic arrays, mostly based on clone collections offering a theoretical density of 1 clone per Mb, versus the high density oligo arrays, that began offering clones

covering thousands of genes, has been overcome by the design and production of genomic arrays displaying the complete tiling-path of the genome in over 30 000 BAC/PAC clones (Barrett *et al*., 2004; Ishkanian *et al*., 2004).

With all this in mind, the choice of a specific platform should be made according to the question that we are trying to answer, biological and practical issues. Are we interested in the characterization of global genomic alteration profiling or are we looking for the detection of small altered regions? Are we interested in the genes within the alteration or in its boundaries? What are the quality and nature of the sample to be analyzed? These generic questions will lead to one or another platform.

3.3.1 Acute myeloid leukemia with normal karyotype

Around 50% of AML cases show myeloid blasts, which are the proliferating cells in this disease, with a normal karyotype. Routine cytogenetic analysis may detect very efficiently genetic aberrations such as chromosome translocations and gains or losses of complete chromosomes, even deletions or duplications can also be detected provided than their size is larger than one chromosome band (medium size, 5 Mb). However, DNA copy number changes encompassing smaller fragments are beyond the scope of microscopic analysis and their identification should be approached by other methods.

In the search for genetic aberrations that may be associated with some of the AML cases with normal karyotype, one reasonable approach is to analyze the presence of DNA copy number changes by array CGH. There is just a single report in the literature where AML with normal karyotype has been studied with array CGH (Raghavan *et al*., 2005). In this study, in which the authors used oligo arrays for detecting loss of heterozygosity, 20% of normal karyotype AMLs were found to have uniparental disomy, a genomic condition not detectable by conventional cytogenetics.

The first question to address for this study was the type of platform to be used. AML is a malignant disease that is frequently characterized by deletions (Cigudosa *et al*., 2003). Of the three types of array CGH approaches BAC arrays seem to be the most reliable to detect losses and low copy number changes, so we decided to use them in our study. In the search for genetic markers in AML, we concentrated on some particular segments of the genome, the subtelomeric regions, which comprise the sequences that are placed immediately after the telomeres towards the centromere. Whereas chromosome telomeres consist of hundreds of repeats of the same sequence (TTAGGG) and their role seems to be to protect chromosome ends through the cell cycle (Blasco, 2005), the genomic regions that immediately follow the repetitive sequences, the subtelomeric regions, contain a high density of genes and segmental duplications (Bailey *et al*., 2002). Within this context, we have been collaborating with other laboratories in preparing a BAC clone collection that completely covered, with overlapped clones, the first megabasepair of the subtelomeric regions, plus a partial coverage of the five following megabases, at a density of one clone per megabase, of all human chromosomes (*Figure 3.2*). We reasoned that studying the subtelomeric regions of AML cases with normal karyotype could provide some useful information regarding the actual genomic status of these otherwise phenotypic abnormal proliferating clones.

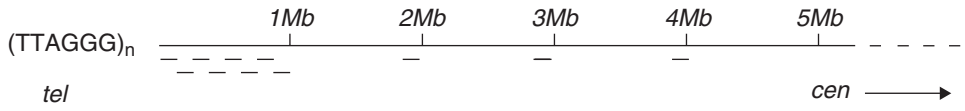

Figure 3.2

Schematic coverage of subtelomeric regions by the BAC clones in the subtelomeric array. The $(TTAGGG)_n$ sequence represents the telomere. Small bars below the line represent BAC clones. The first sequence megabase is completely covered by clones and the following four megabases at 1 clone/Mb. *tel*: telomere; *cen*: centromere.

3.3.2 Cloning of the translocation t(3;8)(p14.1;q24.32)

To clone a translocation point the first approach is to construct a physical map of the breakpoints by placing in order the genomic clones (yeast artificial chromosomes (YACs), BAC, PAC or cosmids) that cover the rearranged chromosomal regions in a comprehensive manner so the region is covered with overlapping clones (contigs). Clones can be easily selected from public databases and obtained for their use as probes for FISH analysis of the aberrant chromosomes. Breakpoint spanning clones can then be identified in a systematic way. However, such FISH investigations typically require several rounds of hybridization starting with clones relatively widely spaced followed by clones at increasingly higher densities until the aberration is defined. This process can be labor intensive and time consuming as many clones have to be hybridized to the patient's chromosomes. The process can be accelerated if we can take advantage of array CGH derived experiments (Fiegler *et al.*, 2003a, 2003b). The complete approach combined array CGH, FISH, polymerase chain reaction (PCR) on flow sorted derivative chromosomes, long-range PCR and sequencing.

Again for this specific experimental design, our choice was a BAC array with a 1 Mb resolution. The choice was due to the availability, robustness, and previously accumulated experience for this kind of experiment.

3.4 Data acquisition

3.4.1 Acute myeloid leukemia with normal karyotype

Material

We selected a series of 16 cases of AML samples with normal karyotype and collected at diagnosis. Array CGH reference DNAs were two separate pools of DNAs (10 female and 10 male) from healthy donors. An AML sample with known chromosome abnormalities as detected by conventional cytogenetic analysis is used to illustrate the resolution power of the technique.

Array CGH platform: subtelomeric array

A specific subtelomeric array CGH platform was constructed in collaboration with Dr Klaas Kok, from the Department of Human Genetics from the

University Hospital of Groningen. This project has been fostered within the European COST B19 Action 'Molecular Cytogenetics of Solid Tumors'. The array included a total of 494 BAC clones that cover 41 subtelomeric regions with a mean of 12 clones/subtelomere. It also contains another 25 control clones from #1 and X chromosomes plus several other clones for quality evaluation of the hybridization. The subtelomeric clone collection was designed to cover the first sequence megabase of each human subtelomeric region in a continuous manner and the following four megabases at a density of one clone per megabase. The array production, that included clone DNA extraction, purification, degenerate oligo priming PCR amplification, as well as printing the DNA onto the slides, was performed essentially as previously described (Westra *et al.*, 2005).

DNA labeling and hybridization

We labeled 1 μg of the sample and reference DNA with Cy3 and Cy5 fluorochromes respectively (Array CGH Protocol 6). Both labeled DNAs were co-hybridized onto the arrays during 40 h. Slides were then washed (Protocol 6, p. 251) and scanned (DNA Microarray Scanner BA, Agilent, Palo Alto, CA, USA) to obtain an image of the signal on each spot of the array. Images were analyzed with GenePix Pro 5.0 (Axon Instruments, Union City, CA, USA), which calculates the signal intensities for each fluorochrome.

3.4.2 Cloning of the translocation t(3;8)(p14.1;q24.32)

In this experimental approach, the array CGH is used as a tool for cloning a translocation breakpoint and the material to be hybridized on the array is DNA obtained from selected chromosome material, specifically those chromosomes that have been rearranged by the translocation event. These experimental procedures have been called *array painting* and they are essentially described by Fiegler *et al.* (2003b).

Material

A lymphoblastoid cell line transformed by Epstein–Barr virus was established from one of the members affected by CC-RCC of the family that carries the translocation. This cell line was the source of chromosomes and DNA for the cloning procedure. To perform the array painting we need to have individualized chromosomes that may be obtained by flow sorting (Carter, 1994). In our experiment, we flow-sorted approximately 500 copies of the rearranged chromosomes #3 and #8, der(3) and der(8), from the patient's cell line and used them as templates for degenerated oligonucleotide priming (DOP)-PCR as described (Telenius *et al.*, 1992). DOP-PCR products from the two derivative chromosomes were differentially labeled with biotin- and digoxigenin-dUTPs by a second round of PCR cycles and hybridized to normal metaphase spreads. Only chromosomal regions comprising the derivative chromosomes showed hybridization signals (*Figure 3.3*).

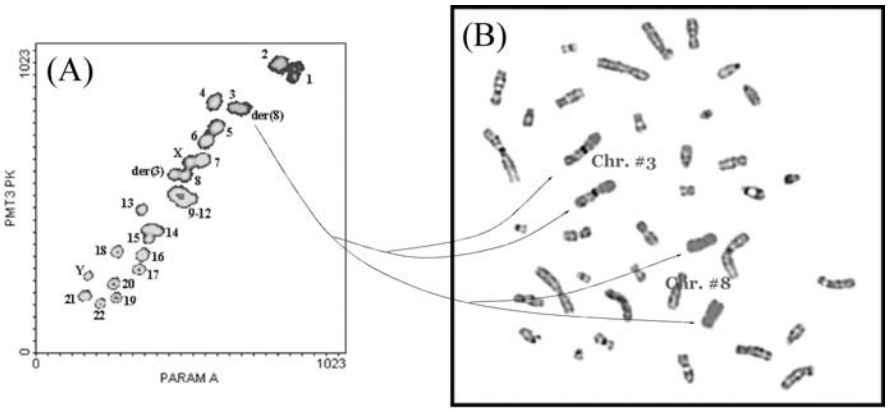

Figure 3.3

(A) Karyotype of the lymphoblastoid cell line from a CC-RCC patient generated by flow sorting. Derivative chromosomes #3 and #8 are indicated. (B) Chromosome painting of the purified der(8) hybridized against a normal metaphase showing that the der(8) is composed of chr. #8 and part of chr. #3. (A color version of this figure is available at the book's website, www.garlandscience.com/9780415378536)

Array CGH platform: genomic array

We used the whole genome array developed at the Sanger Institute (Fiegler *et al*., 2003a) comprising clones selected to be spaced at approximately 1 Mb intervals across the human genome. Only clones corresponding to the sequences present in the sorted chromosomes showed fluorescence and the fluorescence ratio can be either high or low depending on which derivative chromosome the sequence of the clone corresponds to. If a clone on the array spans the breakpoint, sequences from both the derivatives hybridize generating intermediate ratio values.

DNA labeling and hybridization

The two derivative chromosomes were differentially labeled, and hybridized onto the genomic BAC array. Details are given in Protocol 6, p. 249.

3.5 Theory of data analysis

Chromosome aberrations are the basis of developmental abnormalities and cancer because they lead to gains and losses of part of the genome and they include interstitial deletions and duplications, nonreciprocal translocations and gene amplifications. Data output of CGH analysis are ratios between the fluorescence intensity values of test and reference DNAs. The ratio should be 1 when two chromosomal copies of the test and reference DNAs are present in the hybridization reaction. When one of the two copies of a given segment of DNA is lost (heterozygous deletion or mosonomy) the ratio test/reference decreases to 0.5 and it will eventually go to zero where

there is complete loss of both copies (homozygous deletion or nulisomy). One copy gain of a segment (duplication) will produce a three to two ratio (1.5) and to increase the number of copies will raise the ratios following the same scale: four copies, ratio 2; five copies, ratio 2.5; six copies, ratio 3, and so on. Actually, repetitive sequences within the BAC clones and probes, sequence homologies within the genome and normal DNA from stroma or polyclonal nature of the tumor sample reduce the resolution power of the technique, the usual values for one copy gain or loss being 1.4 and 0.6, respectively. Moreover, small variations introduced during the whole labeling and hybridization process generates some dispersion of the data that accounts for ± 0.15 of the expected value. So, normal values can range from 0.85 to 1.15 and clones belonging to altered regions from 1.25 to 1.55 and from 0.45 to 0.75. These ratios are usually expressed as $\log_2$ values.

3.6 Data analysis

The data analyses of both experiments are conducted in the same way. Data ratios between test and reference signals are normalized by print-tip loss with DNMAD (http://bioinfo.cnio.es/, see reference manual for details); this process generates $\log_2$ values. Each clone is spotted in triplicate onto the array and average ratios for these replicas are calculated; if their variation coefficient is higher than 0.2 this average is not calculated and the clone is discarded for the analysis.

Normalized, averaged, clone data ratios are imported into an MS Excel data sheet and related to their precise chromosomal position in the genome. The data is then ordered by this position and plotted (ratio vs. chromosome position). There exist some noncommercial software programs that normalize and/or represent the data in a more graphical view with some added analysis capabilities. Some of these programs are already available on the web, some examples are: CGH-Explorer (at http://www.ifi.uio.no/bioinf/Papers/CGH/); SeeGH (at http://www.bccrc.ca/ArrayCGH); arrayCGHbase (at http://medgen.ugent.be/arrayCGHbase); CAP (on request at bioinfo-cgh@curie.fr) or CGH-Plotter (at http://sigwww.cs.tut.fi/TICSP/CGH-Plotter). In arrays that include well-characterized clones, the clones whose sequence is located on regions or chromosomes with gain/amplification or loss in the test DNA will display ratios that clearly diverge from the normal range ($\log_2$ values: 0.0 ± 0.1). In the AML experiments, the threshold for considering a clone as altered was established for each hybridization as two standard deviations (± 2 SD) of the mean of all clone ratio values. Additionally, there had to be at least two consecutive clones with abnormal ratios to consider a region as altered in its copy number. In the translocation cloning experiment, the sharp transition in ratio values pinpoints the localization of the breakpoint and no threshold values are needed for the analysis. The position of each clone in the genome is known precisely and it allows delimiting of the boundaries of the copy number alteration and also the disclosure of the structure of complex amplicons in which some DNA fragments can be much more amplified than others.

It has to be taken into consideration that DNAs obtained from tumors of polyclonal origin (with different alterations in each cellular clone), or contaminated with normal DNA from surrounding stroma cells, will show

lower resolution in array CGH analysis. It is very important to have this consideration in mind when designing array CGH experiments.

3.7 Summary of the results

3.7.1 Acute myeloid leukemia with normal karyotype

An AML case with a complex karyotype displaying several known chromosome abnormalities was analyzed with the subtelomeric array. The subtelomere hybridization of this control sample confirmed some of the abnormalities expected in these regions as detected by the cytogenetic analysis, allowed the description of new abnormalities that cannot be detected by this method and showed some discrepancies that can be attributed to the inherent technical differences (*Figure 3.4*).

We have looked for DNA copy number changes in the subtelomeric regions of 16 samples of AML patients with normal karyotype at diagnosis. As a previous step, four normal donor DNAs were hybridized against the

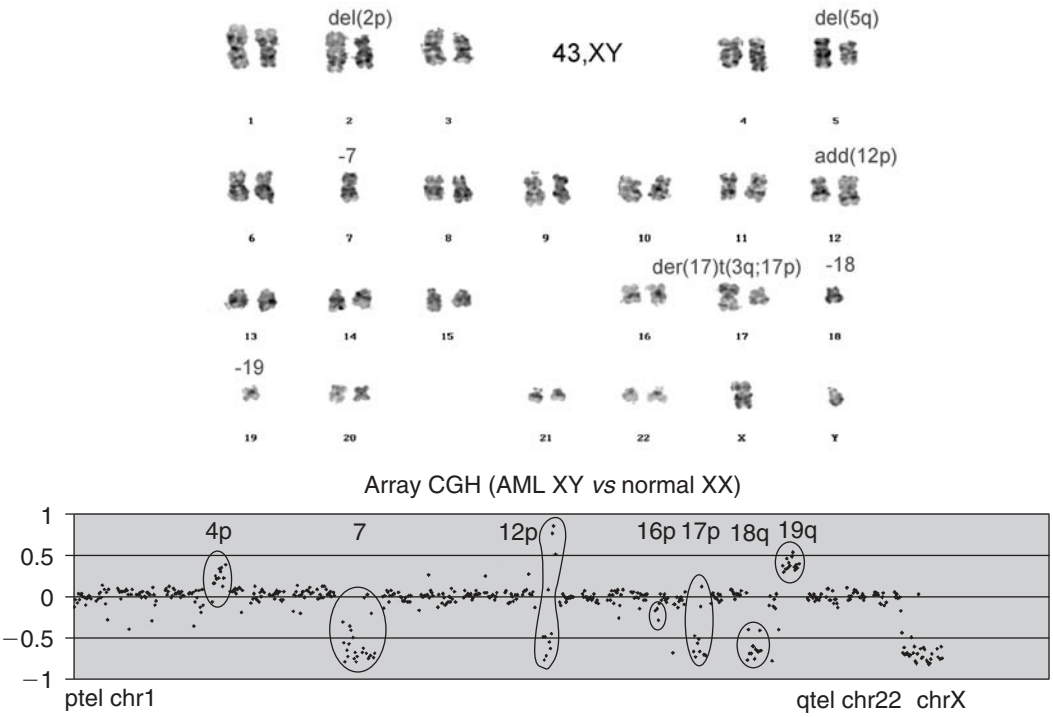

Figure 3.4

G-banded karyotype and subtelomeric array CGH of an AML case with complex karyotype. Major changes are indicated on the G-karyotype and clones localized on duplicated or lost regions are highlighted in the array. Each spot represents a clone in the array, ratio vs. position is plotted. Clones are ordered from the telomere of the *p* arm of chr. #1 to the telomere of the *q* arm of chr. X along the abscissa axis. Array CGH clearly shows complex rearrangements (as in 12p), cryptic gains (4p) and deletions (16p).

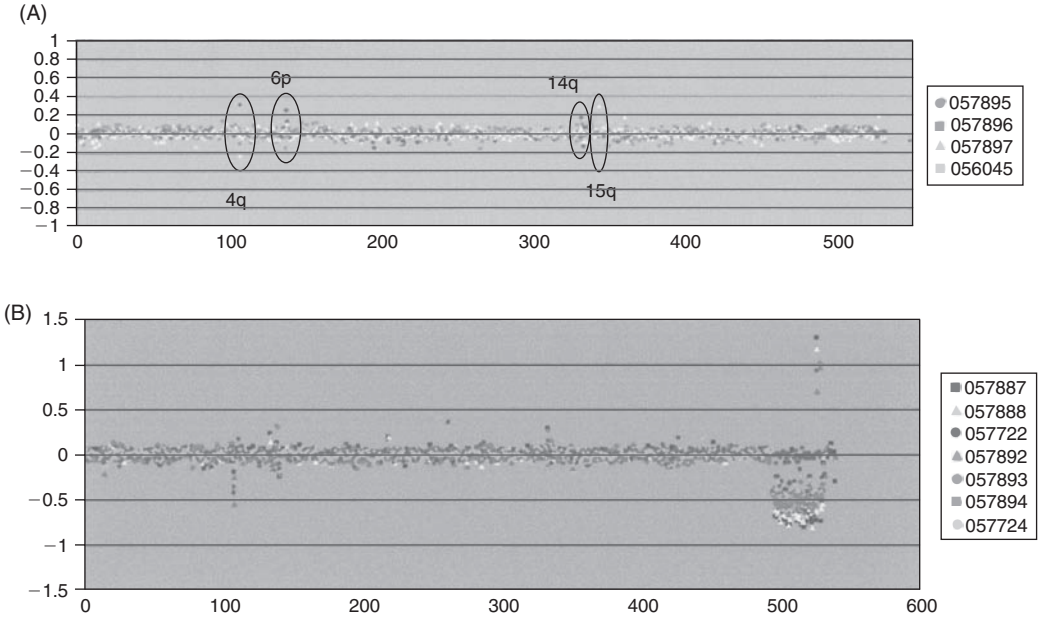

Figure 3.5

Array CGH results for normal and AML cases. Clones are plotted as in *Figure 3.4*. (A) Hybridization results for four normal vs. pooled normal samples. Clones within possible polymorphic regions are indicated. (B) Hybridization results for seven AML cases in which several regions seem to be recurrently altered (including the possible polymorphic ones). In this example, clones belonging to the X and Y chromosomes show ratios clearly different from 0 in samples XY vs. pool XX references (all of them except case #057894). (A color version of this figure is available at the book's website, www.garlandscience.com/9780415378536)

normal pooled reference DNAs. In these control hybridizations four genomic regions, comprising a few BAC clones, showed copy number polymorphisms. These regions were located at the subtelomeres 4q, 6p, 14q and 15q and the polymorphisms were detected as clones that showed abnormal ratios in some of the normal samples but not in others. In the AML cases, these polymorphic regions were also found to be altered in many samples. Thirteen other clones, in 10 different subtelomeric regions, were recurrently detected as abnormal in two to six different cases. These 10 regions were located within the subtelomeres from chromosomes: 4p, 6p, 7p, 15q, 18p, 18q, 19p and 21q (*Figure 3.5*). To investigate if the detected recurrently altered clones are a common feature of AML or a normal polymorphic characteristic of the human population, FISH analysis with probes covering altered regions can be the first option and also the technique of choice for analyzing detected copy number changes in large sets of AML cases.

3.7.2 Cloning of the translocation t(3;8)(p14.1;q24.32)

In our experiment, a sharp transition in ratio values define the translocation breakpoint. The profiles for chromosomes #3 and #8 are shown in

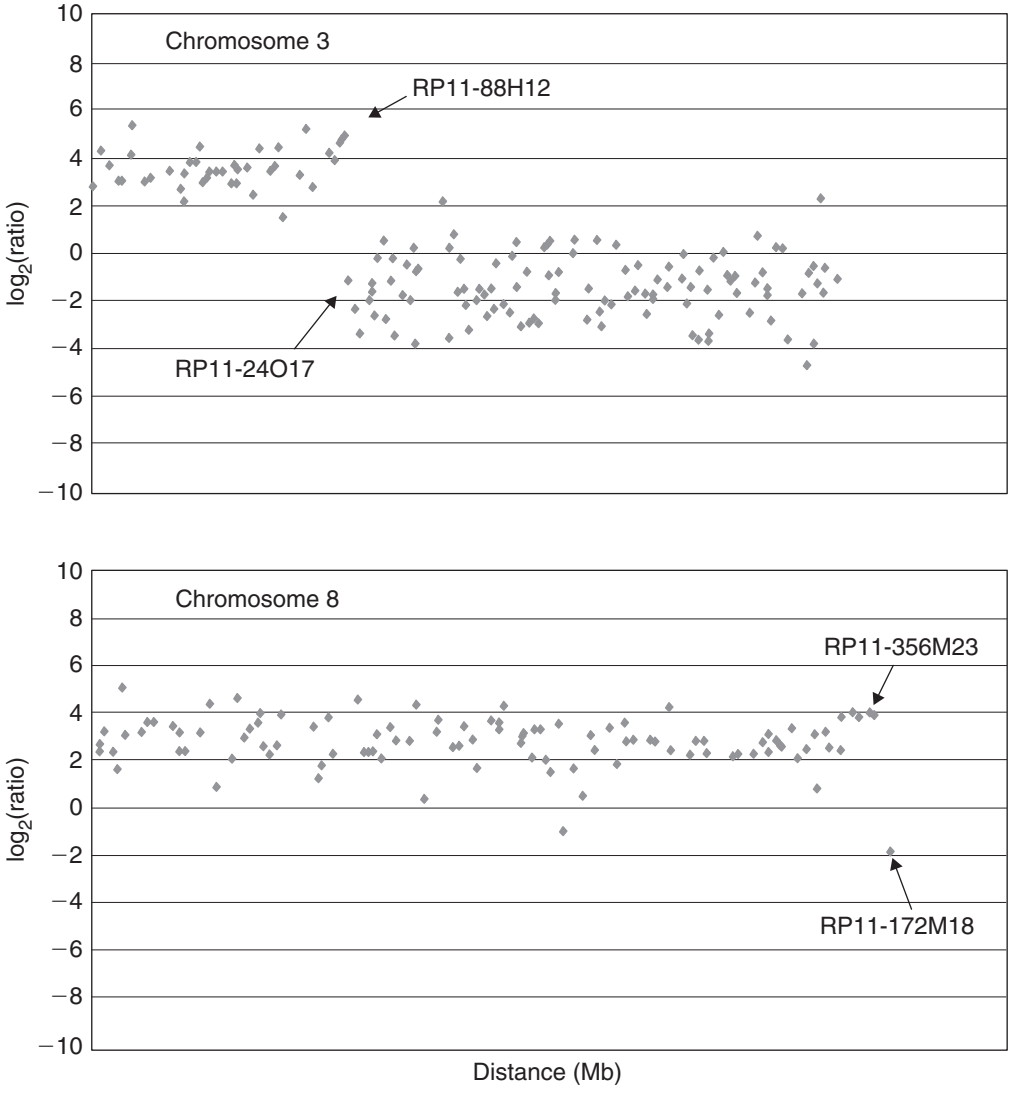

Figure 3.6

Array painting results for chromosomes #3 and #8. Each spot represents a clone in the array. Ratio vs. position is plotted. Clones are ordered from the telomere of the *p* arm to the telomere of the *q* arm along abscissa axes. Sharp ratio transitions between clones localize the translocation breakpoints. Flanking clones are indicated, their exact position in the genome is known.

Figure 3.6. Thus, the chromosome #3 breakpoint lies between clones RP11-88H12 and RP11-24O17; and the clones that defined the chromosome #8 breakpoint are RP11-356H23 and RP11-172M18. After the identification of the BAC clones that span or flank the breakpoints, the cloning process was followed by other molecular approaches.

The complete molecular description of the rearrangement has been essentially reported (Rodriguez-Perales *et al.*, 2004). The analysis of the sequence of the junction of both derivative chromosomes revealed a 5 kb microdeletion at the chromosome #3 breakpoint together with a high density of repetitive motifs and an AT-rich region. No gene had been described at any of the breakpoints and both chromosome #3 and #8 regions flanking the breakpoints were very poor in gene content. Expression analysis of these genes was carried out by RT-PCR but no change was detected in cell lines and patients carrying the translocation.

3.8 Conclusions and suggestions for the general implementation of the case study

In conclusion, the AML study shows how array CGH can uncover genetic/genomic changes that cannot be detected by conventional cytogenetics techniques. For the translocation t(3;8) analysis, the use of the array CGH technique exemplifies how the time consuming FISH mapping procedure to delimit the 1 Mb sequence flanking the translocation site can be reduce to just one assay.

Any biological or medical question that may be caused by genomic copy changes or suspected unbalanced chromosomal rearrangements are suitable to be characterized by array CGH. Translocations, which many times occur with small duplications or deletions in their boundaries, and oncogenes or tumor suppressor genes that can be overexpressed or underexpressed by copy number alterations, are a common feature in tumoral processes. Array CGH is a global genomic technique and allows for global screening and profiling of copy number alterations. This, and the precise localization of the detected changes are the strengths of this kind of array technology. The main effort should be taken in the design of the experiment in relation to the type of array and the resolution needed. Some studies will require specific arrays that cover small chromosomic regions at high clone densities; others, will need whole genome low-density clone configurations. In any case, it will depend on the biological material available and the questions to answer.

Acknowledgments

This work has been partially supported by grants PI040555 and GR/SAL/0219/2004 from Fondo de Investigaciones Santinarias (Instituto de Salud Carlos III) and Comunidad de Madrid, respectively. D.B., S.R.P. and J.C.C. have also been financed by the European COST B19 Action 'Molecular Cytogenetics of Solid Tumours'. Within this Action, we are very grateful to Dr Klaas Kok (Human Genetics Department, University Hospital, Groningen, Netherlands) who has been our main collaborator and helped to develop the subtelomeric array. We have to acknowledge the technical and intellectual support from N.P. Carter from the Wellcome Trust Sanger Institute for the t(3:8) cloning project. We are grateful to Drs M. Urioste and J. Benitez (from the Department of Human Genetics, CNIO) for the t(3;8) translocation patient material and family studies, and to Drs M.J. Calasanz, M.D. Odero, and J. Cervera for providing the AML samples and clinical information.

References

Albertson DG, Pinkel D (2003) Genomic microarrays in human genetic disease and cancer. *Hum Mol Genet* **12(Spec No 2)**: R145–R152.

Bailey JA, Gu Z, Clark RA *et al*. (2002) Recent segmental duplications in the human genome. *Science* **297**: 1003–1007.

Barrett MT, Scheffer A, Ben-Dor A *et al*. (2004) Comparative genomic hybridization using oligonucleotide microarrays and total genomic DNA. *Proc Natl Acad Sci USA* **101**: 17765–17770.

Bignell GR, Huang J, Greshock J *et al*. (2004) High-resolution analysis of DNA copy number using oligonucleotide microarrays. *Genome Res* **14**: 287–295.

Blasco MA (2005) Mice with bad ends: mouse models for the study of telomeres and telomerase in cancer and aging. *EMBO J* **24**: 1095–1103.

Buckley PG, Mantripragada KK, Benetkiewicz M *et al*. (2002) A full-coverage, high-resolution human chromosome 22 genomic microarray for clinical and research applications. *Hum Mol Genet* **11**: 3221–3229.

Carter NP (1994) Bivariate chromosome analysis using a commercial flow cytometer. *Methods Mol Biol* **29**: 187–204.

Carter NP, Vetrie D (2004) Applications of genomic microarrays to explore human chromosome structure and function. *Hum Mol Genet* **13(Spec No 2)**: R297–R302.

Cigudosa JC, Odero MD, Calasanz MJ *et al*. (2003) De novo erythroleukemia chromosome features include multiple rearrangements, with special involvement of chromosomes 11 and 19. *Genes Chromosomes Cancer* **36**: 406–412.

Clark J, Edwards S, Feber A *et al*. (2003) Genome-wide screening for complete genetic loss in prostate cancer by comparative hybridization onto cDNA microarrays. *Oncogene* **22**: 1247–1252.

Fiegler H, Carr P, Douglas EJ *et al*. (2003a) DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. *Genes Chromosomes Cancer* **36**: 361–374.

Fiegler H, Gribble SM, Burford DC *et al*. (2003b) Array painting: a method for the rapid analysis of aberrant chromosomes using DNA microarrays. *J Med Genet* **40**: 664–670.

Hedenfalk I, Ringner M, Ben-Dor A *et al*. (2003) Molecular classification of familial non-BRCA1/BRCA2 breast cancer. *Proc Natl Acad Sci USA* **100**: 2532–2537.

Heim S, Mitelman F (1995) Acute myeloid leukemia. In: *Cancer Cytogenetics* (eds S. Heim and F. Mitelman). Wiley-Liss, New York, pp. 69–140.

Herr A, Grutzmann R, Matthaei A, Artelt J, Schrock E, Rump A, Pilarsky C (2005) High-resolution analysis of chromosomal imbalances using the Affymetrix 10K SNP genotyping chip. *Genomics* **85**: 392–400.

Huang J, Wei W, Zhang J *et al*. (2004) Whole genome DNA copy number changes identified by high density oligonucleotide arrays. *Hum Genomics* **1**: 287–299.

Huntly B, Gilliland DG (2005) Pathobiology of acute myeloid leukemia. In: *Hematology*: *Basic Principles and Practice* (eds R. Hoffman, E.J.J. Benz, S.J. Shattil *et al*.). Elsevier Inc., Philadelphia, pp. 1057–1069.

Hyman E, Kauraniemi P, Hautaniemi S *et al*. (2002) Impact of DNA amplification on gene expression patterns in breast cancer. *Cancer Res* **62**: 6240–6245.

Ishkanian AS, Malloff CA, Watson SK *et al*. (2004) A tiling resolution DNA microarray with complete coverage of the human genome. *Nat Genet* **36**: 299–303.

Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D (1992) Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* **258**: 818–821.

Lindblad-Toh K, Tanenbaum DM, Daly MJ *et al*. (2000) Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. *Nat Biotechnol* **18**: 1001–1005.

Lowenberg B (2001) Prognostic factors in acute myeloid leukaemia. *Best Pract Res Clin Haematol* **14**: 65–75.

Lucito R, West J, Reiner A, Alexander J, Esposito D, Mishra B, Powers S, Norton L, Wigler M (2000) Detecting gene copy number fluctuations in tumor cells by microarray analysis of genomic representations. *Genome Res* **10**: 1726–1736.

Mei R, Galipeau PC, Prass C *et al*. (2000) Genome-wide detection of allelic imbalance using human SNPs and high-density DNA arrays. *Genome Res* **10**: 1126–1137.

Melendez B, Rodriguez-Perales S, Martinez-Delgado B *et al*. (2003) Molecular study of a new family with hereditary renal cell carcinoma and a translocation t(3;8)(p13;q24.1). *Hum Genet* **112**: 178–185.

Monni O, Barlund M, Mousses S *et al*. (2001) Comprehensive copy number and gene expression profiling of the 17q23 amplicon in human breast cancer. *Proc Natl Acad Sci USA* **98**: 5711–5716.

Pinkel D, Segraves R, Sudar D *et al*. (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* **20**: 207–211.

Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, Brown PO (1999) Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* **23**: 41–46.

Pollack JR, Sorlie T, Perou CM *et al*. (2002) Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc Natl Acad Sci USA* **99**: 12963–12968.

Raghavan M, Lillington DM, Skoulakis S, Debernardi S, Chaplin T, Foot NJ, Lister TA, Young BD (2005) Genome-wide single nucleotide polymorphism analysis reveals frequent partial uniparental disomy due to somatic recombination in acute myeloid leukemias. *Cancer Res* **65**: 375–378.

Rauch A, Ruschendorf F, Huang J, Trautmann U, Becker C, Thiel C, Jones KW, Reis A, Nurnberg P (2004) Molecular karyotyping using an SNP array for genomewide genotyping. *J Med Genet* **41**: 916–922.

Rodriguez-Perales S, Melendez B, Gribble SM *et al*. (2004) Cloning of a new familial t(3;8) translocation associated with conventional renal cell carcinoma reveals a 5 kb microdeletion and no gene involved in the rearrangement. *Hum Mol Genet* **13**: 983–990.

Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Benner A, Dohner H, Cremer T, Lichter P (1997) Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* **20**: 399–407.

Telenius H, Pelmear AH, Tunnacliffe A, Carter NP, Behmel A, Ferguson-Smith MA, Nordenskjold M, Pfragner R, Ponder BA (1992) Cytogenetic analysis by chromosome painting using DOP-PCR amplified flow-sorted chromosomes. *Genes Chromosomes Cancer* **4**: 257–263.

Westra JL, Boven LG, van der Vlies P *et al*. (2005) A substantial proportion of microsatellite-unstable colon tumors carry TP53 mutations while not showing chromosomal instability. *Genes Chromosomes Cancer* **43**: 194–201.

Wong KK, Tsang YT, Shen J, Cheng RS, Chang YM, Man TK, Lau CC (2004) Allelic imbalance analysis by high-density single-nucleotide polymorphic allele (SNP) array with whole genome amplified DNA. *Nucleic Acids Res* **32**: e69.