**INTEGRATIVE ONCOGENOMIC ANALYSIS OF MICROARRAY DATA IN HEMATOLOGIC MALIGNANCIES**

Jose A Martínez-Climent[1*], Lorena Fontan[1], Vicente Fresquet[1], Eloy Robles[1], María Ortiz[2], Angel Rubio[2]

(1) Center for Applied Medical Research, Division of Oncology, University of Navarra, Pamplona, Spain;
(2) CEIT and TECNUN, University of Navarra, San Sebastián, Spain

(*) Corresponding author: jamcliment@unav.es

## SUMMARY

**INTEGRATIVE ONCOGENOMIC ANALYSIS OF MICROARRAYS DATA IN HEMATOLOGIC MALIGNANCIES**


## Abstract

During the last decade, gene expression microarrays and array-based comparative genomic hybridization (array-CGH) have unraveled the complexity of human tumor genomes more precisely and comprehensively than ever before. More recently, the simultaneous assessment of global changes in mRNA expression and in DNA copy number through “integrative oncogenomic” analyses has allowed researchers the access to results uncovered through the analysis of one-dimensional datasets, thus accelarating cancer gene discovery. In this chapter, we discuss the major contributions of DNA microarrays to the study of hematological malignancies, focusing on the integrative oncogenomic approaches that correlate genomic and transcriptomic data. We also present the basic aspects of these methodologies and their present and future application in clinical oncology.

## 1. INTRODUCTION.

The application of gene expression microarrays has allowed the definition of common patterns of gene expression that can distinguish pathologically different tumors, but has also revealed degrees of heterogeneity between and within tumors.[1-5] Alizadeh and colleagues were the first to use microarrays to identify subtypes of a single disease (diffuse large B cell lymphoma) that could only be defined by their gene expression patterns.[6] Since this pioneering report, gene expression profiling has made significant contributions in basic and applied cancer research by providing useful prognostic biomarkers, defining novel oncogenic pathways, characterizing molecular portraits of transformation and metastasis, and revealing unique gene signatures of therapeutic response.[5,7-9] Today, the extrapolation of some of these microarray findings to more efficient and cost-effective laboratory techniques, such as quantitative PCR, flow cytometry or immunohistochemistry, are remarkable examples that place basic science closer to clinical medicine.[8] A different DNA microarray technology, termed comparative genomic hybridization (CGH) to microarrays (from now on, array-CGH), can detect and map changes in the DNA copy number that are present in tumors but not in the corresponding non-tumoral germline sequences.[10-12] In seminal papers, Pinkel, Lichter and colleagues used high-resolution array-CGH with bacterial artificial chromosomes (BACs) as clones to precisely define amplicon structures and deletion borders in tumors, mapping the corresponding gene loci targeted by the amplification and deletion processes.[12-14] Since then, notable improvements in the resolution and sensitivity of current genome-wide array-CGH platforms have made possible the accurate screenning for genome-wide aberrations in large tumor sets.[11,15,16] More recent oligonucleotide-based single nucleotide polymorphism (SNP) microarrays were able to detect not only DNA copy number changes but also copy-neutral genetic aberrations such as loss of heterozygosity caused by uniparental disomy.[17] Overall, systematic scanning of cancer genomes using array-CGH has served to describe patterns of genetic alterations linked to the genesis and dissemination of human tumors.

Yet gene expression microarrays and array-CGH are both mature technologies, the simultaneous assessment of global changes in mRNA expression and in DNA profile through "integrative oncogenomics" represents a relatively novel approach that attempts to accelarate cancer genome annotation and gene target discovery at a genome scale.[5,18] Using these comparative systems, which need support from robust bioinformatic tools, researchers are having access to results uncovered through the analysis of one-dimensional datasets. Notable examples are the identification of specific chemotherapy response signatures by microarray analyses of multiple human biopsies and human-like tumors from genetically manipulated mice, or the construction of regulatory genetic networks where participating cancer genes can be functionally characterized in proper molecular and cellular contexts.[18] In this chapter, we discuss the major contributions of DNA microarrays to the study of hematological malignancies, focusing on the integrative oncogenomic approaches that correlate genomic and transcriptomic data. We also present the basic aspects of these methodologies and their application in clinical oncology.

## 2. APPLICATION OF GENE EXPRESSION MICROARRAYS IN HEMATOLOGICAL TUMORS

The study of hematological malignancies has particularly benefited from gene expression analysis, and crucial discoveries about diagnosis, prognosis and pathogenetic mechanisms of these diseases have been made. In acute myeloid leukemia (AML), good prognostic subgroups are defined by the presence of specific chromosomal rearrangements such as the translocations t(8;21) and t(15;17) or the inversion of chromosome 16, whereas translocations afecting the *MLL* gene in chromosome 11q23 or deletions of chromosome 5q or 7q characterize poor prognostic subgroups.[19,20] Gene expression profiling has been able to identify these leukemia subgroups with high accuracy. In two landmark papers, gene expression profiling detected not only previously defined genetically and prognostically subgroups in AML but also novel clusters with adverse prognosis.[21,22] A different approach evaluated the expression profiling of CD34+ hematopoietic stem/progenitor cells, revealing distinct subtypes of therapy-related AML.[23] In B-cell acute lymphoblastic leukemia (B-ALL), gene expression

profiles distinguished each of the prognostically important leukemia subtypes, including those with specific chromosomal translocations: t(1;19)-*E2A-PBX1*, t(9;22)-*BCR-ABL*, t(11q23)-*MLL* and t(12;21)-*TEL-AML1*, as well as those with hyperdiploidy with >50 chromosomes. Further, within some of these genetic subgroups, those patients who eventually relapsed presented typical gene expression patterns that allowed their recognition.[24] Gene expression studies have also deciphered novel oncogenic pathways in childhood T-cell acute lymphoblastic leukemia (T-ALL),[25] and in adult T-cell lymphoma in leukemic phase.[26] Notably, these studies identified T-ALL subgroups with molecular signatures associated with favorable prognosis (*HOX11*), while those expressing *TAL1*, *LYL1* or *HOX11L2* presented much worse responses to treatment.[25] A common aspect of these studies is the identification of molecular subgroups defined by oncogenes that are aberrantly expressed in the absence of chromosomal abnormalities. Therefore, gene expression microarrays can identify all leukemias within identical molecular subgroups, including cases with typical chromosomal rearrangements but also others that would be missed by standard cytogenetic and molecular techniques. One additional goal of gene expression microarrays has been to search for therapeutic targets in patients with leukemia. *FLT3* mutations, a common genetic abnorality in AML, is an independent prognostic indicator of poor outcome and response to standard chemotherapy.[27] In a cDNA microarray study of childhood leukemias, *FLT3* was found to be overexpressed in patients carrying *MLL* gene translocations. Subsequent studies showed that FLT3 inhibitors are active against leukemias with *MLL* rearangements *in vitro* and *in vivo*.[28,29] In conclusion, it is unquestionable that gene expression arrays have had an enormous impact in our current understanding of acute leukemias. To move on, clinical trials should evaluate novel therapies in patients who are stratified according to the molecular profiles determined at the time of diagnosis.

Molecular profiling has also been crucial in deciphering the pathogenesis of B-cell malignancies. Diffuse large B cell lymphoma (DLBCL) can be divided into molecular subgroups based on their cellular origins that significantly differed in therapy response and cure rate.[6,30] Importantly, these molecular subsets of disease, namely germinal center DLBCL (GC-DLBCL) and activated B-cell

DLBCL (ABC-DLBCL), were only distinguishable by gene expression profililng and not by other current diagnostic methods.[6] Following these studies, immunohistochemistry-based assays were developed to classify ABC-DLBCL and GC-DLBCL cases on a routine basis.[31] Additionally, the lymphochip survival-prediction data were further validated through measuring single expression of six genes (*LMO2*, *BCL6*, *FN1*, *CCND2*, *SCYA3*, and *BCL2*) by quantitative PCR, which was sufficient to predict overall survival in patients with DLBCL treated either with CHOP or more recently, with CHOP plus the anti-CD20 monoclonal antibody rituximab (R-CHOP).[32,33] Most importantly, translation of this molecular knowledge to the clinic may lead to therapeutic tailoring in patients with DLCBL. For instance, ABC-DLBCL cases show constitutive activation of NF-kB genes and thus respond uniquely to NF-kB inhibitors.[34] Further refinement of these investigations showed a cooperative signaling through the STAT3 and NF-kB pathways in a subset of ABC-DLBCL cases. A small-molecule inhibitor of JAK signaling, which blocked STAT3 signature expression, was toxic only for ABC-DLBCL lines and synergized with an NF-kB inhibitor.[35]

Additional Affymetrix microchips ptofiled molecular signatures of DLBCL cases with different responses to standard chemotherapy, thus revealing unique pathways associated with poor responses.[36] Moreover, the use of multitple clustering and gene set enrichment analysis allowed the identification of three discrete subsets of DLBCL termed "oxidative phosphorylation," "B-cell receptor/proliferation," and "host response" (HR), pointing out that the tumor microenvironment and the host inflammatory response are defining features in DLBCL.[37] These molecular routes altered in subsets of patients with DLBCL may represent important targets for therapeutic intervention using specific drugs,[37-40] some of which are being used in phase II clinical trials.[41,42] Subsequent studies have revealed a transcriptional signature with differential expression of BCL6 target genes that can accurately identify DLBCL cases carrying genetic alterations of the *BCL6* oncogene. Notably, the DLBCL subgroup with the BCL6 expression signature is uniquely sensitive to BCL6 inhibitors.[43-45]

Other remarkable achievements of gene expression analysis of B-cell malignancies include the recognition of molecular links between apparently

different entities such as Hodgkin disese and primary mediastinal B-cell lymphoma, suggesting a putative common cellular origin;[46-48] the discovery of better prognostic markers, such as ZAP70 expression measurement as an indicator of survival length of patients with B-cell chronic lymphocytic leukemia (B-CLL);[49-52] a more accurate molecular definition of Burkitt lymphoma that expands the spectrum of the WHO criteria for this disease;[53,54] the correlation of survival duration in patients with follicular lymphoma with gene expression profiles reflecting an interaction between tumor cells and infiltrating immune cells;[55,56] the understanding of the roles of cell cycle control and DNA repair pathways that correlate with cell proliferation and clinical outcome in mantle cell lymphoma;[57,58] the definition of a typical gene expression profiling of hairy cell leukemia that reveals a phenotype related to memory B cells with altered expression of chemokine and adhesion receptors;[59] and the investigation of the multistep transformation of monoclonal gammopathy of undetermined significance to multiple myeloma by global gene expression analysis.[60-63]

Recently, expression microarrays have been used to investigate questions on leukemia biology and therapy in more complex functional model systems. Krivtsov and colleagues demonstrated in a mouse model of leukemia initiated by *MLL-AF9* oncogenic fusion that leukemia stem cells maintain the global identity of the progenitor cells from which they arosed while activating a limited stem-cell or self-renewal-associated gene expression program characteristic of hematopoietic stem cells.[64] In a different report, Ngo and colleagues used a doxycycline-inducible retroviral vector for the expression of small hairpin RNAs (shRNAs) for 2,500 human genes in ABC-DLBCL and GC-DLBCL cell lines. Each vector was engineered to contain a unique 60-base-pair 'bar code', allowing the abundance of an individual shRNA vector within a population of transduced cells to be measured using microarrays of the bar-code sequences. Results determined that a subset of shRNA vectors was depleted from the transduced cells when shRNA expression was induced, uncovering *CARD11* gene as a key component responsible for the constitutive NF-kB activation in ABC-DLBCL but not in GC-DLBCL.[65] Further validating the screening, this group of investigators found mutations in exons encoding the coiled-coil domain of *CARD11* gene that activated NF-kB pathway in 9,6% of patients with ABC-DLBCL.[66] These data

point out CARD11 as an atractive therapeutic target in this lymphoma subgroup, but we still do not know the genetic lessions in the remaining 90% of ABC-DLBCL cases with NF-kB signaling. Similar massive genomic screenings using RNAi have been recently applied to other tumor cell types, providing alternative functional information of cancer cells.[67-69] Palomero and colleagues applied gene expression microarrays to T-ALL cell lines that were classified as sensitive or resistant to gamma-secretase inhibitors, which block a proteolytic cleavage required for NOTCH1 activation. Among the genes targets that were found differentially expressed was the tumor suppressor *PTEN*. Further investigations demonstrated that NOTCH1 regulates the expression of PTEN and the activity of the phosphoinositol-3 kinase (PI3K)-AKT signaling pathway in normal and leukemic T cells.[70] This novel observation suggests the need to simultaneously inhibit both pathways to improve therapeutic efficacy in T-ALL.

Microarray technologies have been also used to measure global expression of a class of small non-coding RNA species, known as microRNAs in tumors.[71,72] One of the cancers where miRNA profiling has provided critical information is B-CLL. Patients with B-CLL and prolongued survival were characterized by down-regulation of miR-15a and miR-16-1 located at 13q14.3.[73,74] On the other hand, B-CLL cases with unmutated $IgV_H$ or with elevated expression of ZAP70 showed high levels of *TCL1* due to low-level expression of miR-29 and miR-181, which directly target this oncogene.[75] These data suggest that B-CLL is a disease in which the main pathogenetic alterations may occur in miRNAs.

## 3. DNA COPY NUMBER VARIATION IN LEUKEMIA, LYMPHOMA AND MYELOMA

Tumor genomes usually show a large diversity of abnormalities, ranging from point mutations to overt chromosomal aberrations, that have been accumulated through the process of malignant transformation. Albeit array-CGH cannot detect small sequence mutations or reciprocal chromosomal translocations, this technology has facilitated the description of global portraits of DNA copy number aberrations in tumors with high precision and resolution.[76] In hematologic malignancies, array-CGH has led to the identification of known amplicons as well

as hidden gene amplifications unappreciated by other genetic screens, such as those containing c-*MYC* in 8q24 in DLBCL and follicular lymphoma, *REL* ad *BCL11A* genes in DLBCL, *JAK2* and *PDL2* in primary mediastinal B cell lymphoma, *BMI1* in mantle cell lymphoma, and *CCND3* and *BYSL* in DLBCL.[47,77-80] Some of these aberrations showed unexpected patterns of complexity only discovered through array-CGH studies. For instance, a detailed mapping of the 18q21.3 amplicon disclosed two major target sites involving *BCL2* and *MALT1* genes in follicular lymphoma and MALT lymphoma, respectively.[81] A number of the amplicon targets included genes commonly involved in *IG*-related chromosomal translocations.[82] Moreover, some of these genomic amplifications seem to originate after regional chromosomal translocations. For example, chromosome 8q24 sequences surrounding *c-MYC* loci are frequently amplified in many B-cell malignancies carrying a previous t(8;14)(q24;q32).[79,83] This is also the case for the lymphomas develped in deficient mice for *P53* and nonhomologous end-joining (*NHEJ*) genes, that present complex genomic rearrengements with coamplification of *c-MYC* (chromosome 15) and *IgH* (chromosome 12) sequences.[84] Whether these secondary genomic amplifications have any biological impact or merely reflects local genomic instability remains unknown. Overall, genomic amplification functions as a mechanism of oncogene activation alternative to *IG*-related translocations in B-cell lymphoma (Figure 1A). In one interesting study, amplification of 7p22 in adult T-cell leukemia/lymphoma pinpointed *CARD11* as the possible target, a gene that has been recently implicated as a NF-kB activator in ABC-DLBCL.[66,85] However, its implication in T-cell leukemia/lymphoma, which also shows constitutive NF-kB activation, awaits further studies. Non-coding microRNAs may also be te target of genomic amplification. In the chromosome 13q31.3 amplificon, commonly observed in follicular lymphoma, GC-DLBCL, mantle cell lymphoma and splenic marginal zone lymphoma, the miR-17-92 cluster results in overexpression of up to 500 fold (Figure 1B). These microRNAs positively target *c-MYC* oncogene in B-cell lymphomas,[86,87] and have been implicated in several biological processes depending on the cellular context, such as the control of monocytopoiesis through AML1 targeting and M-CSF receptor upregulation,[88] or the promotion of proliferation and the inhibition of differentiation of lung epithelial progenitor cells.[89] In three recent studies, the miR-17-92 cluster has been shown

to have critical roles in normal B-cell lymphopoiesis as well as in the pathogenesis of B-cell lymphomas and some autoinmune disorders in mice.[90-92]

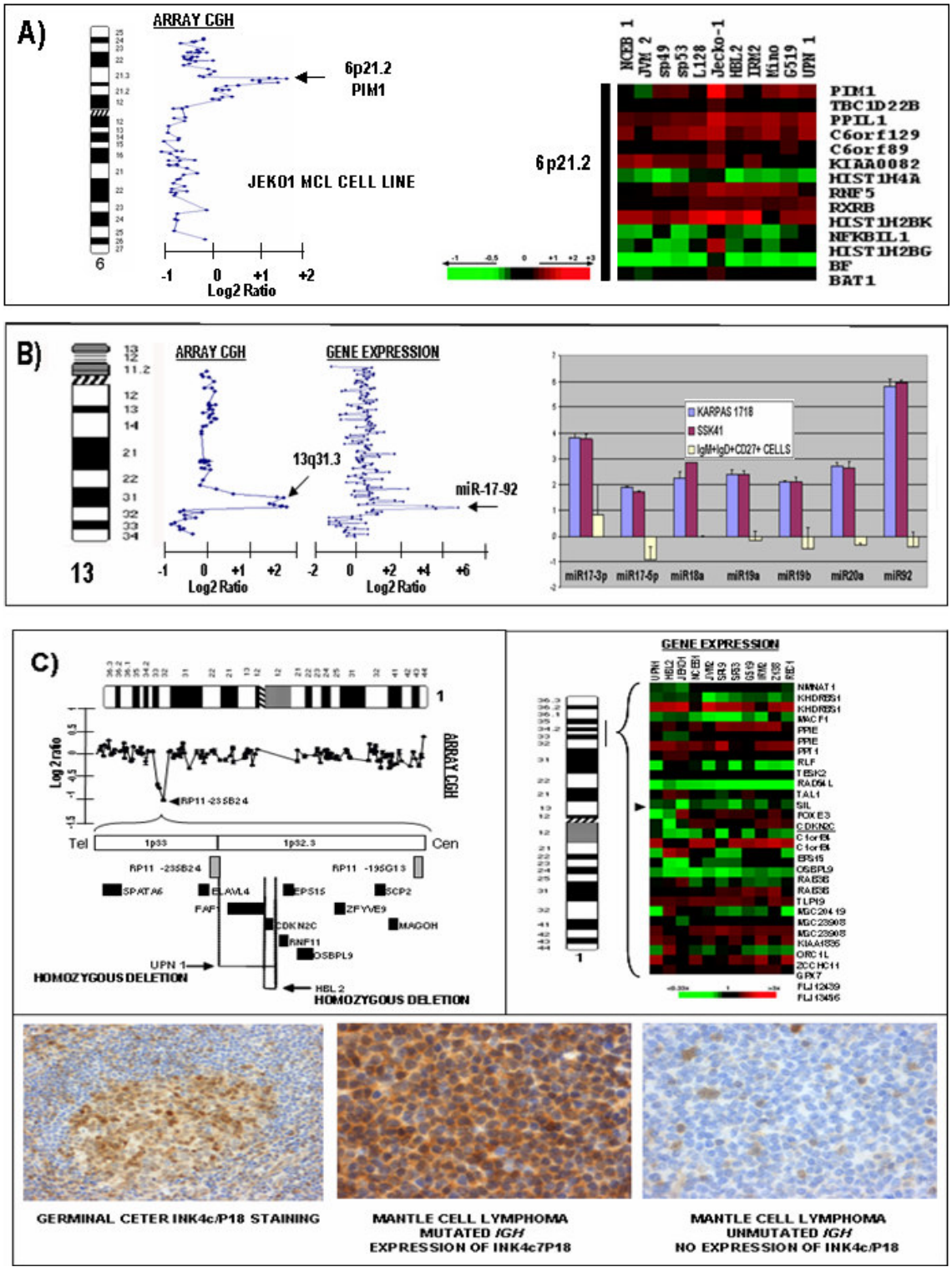


Figure 1. A) Array CGH representation showing genomic amplification of 6p21.2 (left) leading to overexpression of *PIM1* oncogene in mantle cell lymphoma (right). B) High-level amplification of 13q31.3 targeting the miR-17-92 cluster in Karpas 1718 cell line (left). Using quantiative PCR, overexpression of the 7 miRs in the cluster is demostrated in comparisson to normal marginal zone-derived IgM+IgD+CD27+ B cells (right). C) Homozygous deletion of *INK4c/P18* gene in mantle cell lymphoma detected by array CGH (upper left). The genomic loss is associated with low expression at the RNA level (shown in gene expression microarrays, upper right) and at the protein level (shown by immunohistochemistry (bottom).

Localization and rapid delineation of areas of genomic loss has been one of the major achievemnts of the application of array-CGH in tumors. Initial reports mapped deletions of known tumor suppressor genes such as *P53*, *P16*, *ARF* and *FHIT* that resulted inactivated with variable frequencies in many cancer types.[93-99] Similar studies have also discovered the loci of novel genes inactivated in B-cell malignancies, such as the pro-apoptotic *BIM* in chromosome 2q13 that is disrupted by minimal bi-allelic deletions in mantle cell lymphoma; the *PRMD1* gene in chromosome 6q21, frequently targeted by deletion of one allele and by truncatig mutation of the remaining allele in ABC-DLBCL but not in GC-DLBCL; the NF-kB inactivator gene *TNFAIP3*, which shows bi-allelic deletion in ocular MALT lymphoma, follicular lymphoma and DLBCL, and the *INK4c/P18* gene targeted by bi-allelic deletion or heterozygous loss and mutation of the remaining allele in mantle cell lymphoma (Figure 1C).[100-108] The use of DNA microarrays containing 32,000 overlapping BACs has enhanced the detection sensitivity of array-CGH devices, for instance by mapping intra immunoglobulin gene deletions at 2p11 and 22q11 chromosomes as small as 130 kb.[109] Positional identification of suppressor genes in genomic deletions has also been investigated by array-CGH in mouse models of cancer. A pioneering report by Hodgson and colleagues used array-CGH to scan the genomes of mouse islet carcinomas, revealing regional alterations that are syntenic to human genome sequences containing candidate oncogenes and suppressor genes.[110] In a different study, Mao and colleagues studied radiation-induced lymphomas from P53 deficient mice. Lymphomas from P53+/- mice, but not those from P53-/- mice, showed frequent loss of heterozygosity and a 10% mutation rate of *FBXW7/hCDC4*, a gene encoding a ubiquitin ligase implicated in the control of chromosome stability. Further investigations showed that FBXW7+/- mice have greater susceptibility to radiation-induced tumorigenesis, but most tumours retain and express the wild-type allele, indicating that *FBXW7* is a haploinsufficient tumour suppressor gene.[111] However, interpretation of these microarray approaches in mouse models of cancer may be difficult, because BAC array-CGH studies of normal genomes from different strains of common laboratory mice revealed important segmental and sequence variations.[112]

Genome-wide approaches using Affymetrix SNP-CGH technology have recently demonstrated their power of high-resolution to identify new molecular lesions in different cancer types. These sudies revealed deletion, amplification, intragenic mutation and structural rearrangement in genes encoding principal regulators of B lymphocyte development and differentiation in 40% of B-ALL cases, preferentially targeting the *PAX5* gene.[113] In accordance with these unexpected data, Cobaleda and colleagues reported that mice lacking PAX5 in mature B cells developed aggressive B-cell malignancies, which were identified by their gene expression profile as progenitor cell tumours.[114] In addition, CCAAT enhancer-binding protein (CEBP) transcription factors CEBPA and CEBPB, which down-regulate PAX5 in B-cells,[115] have been shown to be over-expressed in B-ALL through *IG*-related translocation.[116] Collectively, these findings suggest that block of genes controlling B-cell development and differentiation, especially of PAX5 either by genetic inactivation or by functional suppresion through CEBP family members, contributes to B-cell ALL pathogenesis.

Efforts are being made to translate the correlation of genomic and clinical data detected by DNA microarrays into more feasible and applicable clinical tests. B-CLL can be considered as a prototype disease where length of patient survival can be predicted by the presence of chromosomal aberrations associated with poor prognosis (deletions of chromosomes 11q22-q23 and 17p13) or with favorable prognosis (deletion of 13q14 or cases with normal karyotype).[117] Routine detection of these genetic alterations is currently performed by FISH technique of bone marow or peripheral blood samples obtained at diagnosis in many laboratories throughout the world.[118] An automated BAC array-CGH detected with high precission all these changes in a series of 106 patients with B-CLL, especially in those with >50% of tumoral cells in blood or marrow samples.[119] Mantle cell lymphoma (MCL) is also characterized by a set of genomic aberrations that target genes involved in the pathogenesis of the disease. Examples include the genomic amplification of chromosomes 8q24 affecting *c-MYC*, 10p13 involving *BMI1* oncogene and 11q13 targeting *CCND1*/cyclin D1, and the losses of 8p21.3 including *TRAIL-R1/R2* genes, 9p21 (*INK4A/ARF*), 11q23 (*ATM*) and 17p13.1 (*P53*).[83,120,121] The pattern of these alterations has been correlated with tumor phenotypes and thus, blastoid variants

of MCL usually display inactivation of *P16/INK4A* and *P53* genes whereas indolent forms of MCL, usually having mutated $IgV_H$ genes, frequently show deletion of chromosome 8p.[83,120,121] Although most patients with MCL show poor clinical outcome with current immunochemotherapy regimens, the long-term survivors can be identified by a characteristic genomic profile defined by the absence of deletions of *P53*, *P16/ARF* and chromosome 9q21-q22, and by the presence of the deletion of chromosome 1p21-p22.[83,122] In both B-CLL and MCL, the development of disease-specific CGH microchips may be of value in the clinic, as should allow testing the genomic profiles as prognostic and predictive factors of response to novel therapies. In a recent report, high-density SNP-CGH arrays were used to analyze genome-wide changes of copy number and allele status in B-CLL samples from patients that were sensitive or resistant to MDM2 inhibitors. These studies conclusively demonstrate that *P53* status is the major determinant of response to MDM2 inhibitors in B-CLL.[123] In a study of 107 follicular lymphoma (FCL) diagnostic biopsies with an array CGH platform containing over 26,819 BAC clones covering >95% of the human genome, 68 regional alterations were identified in >10% of cases. Importantly, 11 of these areas were independent predictors of overall survival using a multivariate analysis that included IPI score. Further, two of the 11 regions (deletions of 1p36 and 6q21-q24) were also predictors of transformation risk (Cheung et al, in press). These genetic data may be useful to identify FCL high-risk patients as candidates for risk-adapted therapies.

The acquisition of uniparental disomy (UPD) is a common event in cancer. Genome-wide SNP analysis has revealed large-scale cryptic regions of UPD in many hematologic tumors. In AML, this alerations are non-random and contain homozygous mutations in genes known to be mutational targets in leukemia (*WT1*, *FLT3*, *CEBPA*, and *RUNX1*).[124,125] A high proportion of patients with myeloproliferative disorders, including polycythemia vera, essential thrombocythemia and chronic idiopathic myelofibrosis, carry a dominant gain-of-function mutation of *JAK2.*[126-128] Using SNP-CGH, the UPD of chromosome 9p typical of these entities has provided the molecular mechanism of homozygos mutation of *JAK2* in these entities.[129] These data imply that mutation of one allele precedes mitotic recombination, which acts as a "second hit" responsible for

removal of the remaining wild-type allele which is substited by a copy of the mutated allele. Additional examples are the identification of UPD surrounding the *NF1* gene locus in cases of juvenile myelomonocytic leukemia associated with neurofibromatosis.[130] In lymphoma, the mutation status of genes within areas of UPD is less established, although bi-allelic mutations of *P53* and *P16/ARF* have been reported in cases with UPD of 17p and 9p, respectively. SNP-CGH arrays described that in mantle cell lymphoma and in follicular lymphoma, most areas of UPD were coincident with known regions of chromosome deletion.[104,105,131] However, UPD was also observed in chromosome 6p in 20-30% of initial biopsies from patients with follicular lymphoma, an area not usually targeted by DNA copy number changes. To date, the gene or genes involved in this area have not been yet detected.[104,105]

A different application of high-density SNP-CGH arrays has been the genome-wide linkage search of 206 families with B-CLL. These studies indentified potential susceptibility loci on chromosomes 2q21.2, 6p22.1 and 18q21.1. Notably, none of the regions coincided with areas of common chromosomal abnormalities frequently observed for B-CLL.[132] These findings strengthen the argument for an inherited predisposition to B-CLL that might explain familial aggregation, and support similar microarray studies in other familial cancers with unknown causing genes.

## 4. INTEGRATIVE ONCOGENOMICS AS A TOOL TO DISCOVER NOVEL CANCER GENES

Initial comparative genomic studies evaluated the degree to which DNA copy number alterations contributes to variations in the transcriptional program of tumors.[133] Using cDNA microarrays, Pollack and colleagues found that 62% of highly amplified genes in breast tumors showed moderately or highly elevated expression. However, the influence of low-level DNA copy number changes was much more limited and only 12% of all the variation in gene expression among the breast tumors was directly attributable to underlying genomic dosage.[134] Using again beast cancer as a model disease, Hyman and colleagues reported that both high- and low-level copy number changes had a substantial impact on

gene expression, with 44% of the highly amplified genes showing overexpression and 10.5% of the highly overexpressed genes being amplified.[135] A third study focused on the process of transformation of follicular lymphoma (FCL) to DLBCL, which is observed in over one-third of patients with FCL and is generally characterized by an aggressive clinical course and refractoriness to treatment. Parallel array-CGH and gene expression analyses revealed that FCL transformation was accompanied with a variable spectrum of recurrent genomic imbalances and gene expression changes. Among the ~600 genes that presented deregulated expression in the transformation phase, up to one-third showed correlation with DNA copy number variation.[136] Overall, these reports concluded that a fraction of transcriptomic modifications are consequence of genomic changes in tumors.

Since these studies, more sophisticated bioinformatic methods were developed for determinig if altered patterns of gene expression correlate with chromosomal abnormalities. One of these softwares is ChARM (Chromosomal Aberration Region Miner), a robust and accurate expectation-maximization based method for identification of segmental aneuploidies from gene expression and array CGH microarray data, sensitive enough to detect statistically significant and biologically relevant subtle changes in mixed populations of cells.[137] Likewise, DIGMAP is a powerful computational tool enabling the coupled analysis of microarray data with genome location.[138] More complex devices include VAMP softwares (Visualization and Analysis of array-CGH, transcriptome and other Molecular Profiles), developed as a graphical user interface for visualization of CGH arrays, transcriptome arrays, SNP-CGH arrays, loss of heterozygosity results (LOH), and Chromatin ImmunoPrecipitation arrays. The interface offers the possibility of looking for recurrent regions of alterations, confrontation to transcriptome data or clinical information, and clustering.[139] ARACNE is a different algorithm designed to scale up to the complexity of cellular regulatory networks present in microarray profiles, based on a theoretic approach that eliminates indirect interactions inferred by co-expression methods. For instance, authors demonstrated and validated a complex interactive network among the transcriptional targets of the *c-MYC* oncogene in B-cell lymphomas.[140]

One of the major advances of integrative oncogenomic approaches has been the identification of novel cancer genes. In one landmark report, Garraway and colleagues identified *MITF* (microphthalmia-associated transcription factor) as the target gene of a melanoma amplification by integrating SNP-CGH array maps with gene expression signatures derived from the NCI60 cell lines. Further investigation demonstrated that *MITF* represents a 'lineage survival' oncogene required for both melanoma development and metastatic spread.[141] In the study by Yu and colleagues, the power of integrating multiple diverse genomic data of prostate cancer models (*in vitro* cell line, *in vivo* tumor profiling, and genome-wide location data) to search for key targets genes of the Polycomb family protein EZH2, showed *ADRB2* gene as a critical mediator of beta-adrenergic signaling.[142] A number of additional papers have applied similar genetic screens to mouse models of cancer to discover new oncogenes. In a screen for gene copy-number changes in mouse mammary tumors, a 350-kb amplicon from a region syntenic to a locus amplified in human cancers at chromosome 11q22 was detected. This amplicon contained only one gene, *YAP*, which encodes the mammalian ortholog of Drosophila Yorkie (Yki), resulted a regulator of cellular proliferation and apoptosis in epithelial cells.[143] In a mouse model of hepatocarcinoma, genome-wide analyses of tumors revealed a similar amplification at mouse chromosome 9qA1 syntenic to human chromosome 11q22. Gene-expression analyses delineated *cIAP1* and *YAP* as candidate oncogenes that cooperated to promote tumorigenesis.[144] A different study characterized metastatic variants in an induced mouse model of melanoma, identifying an acquired focal chromosomal amplification that corresponded to a much larger amplification in chromosome 6p25 in human metastatic melanomas. Further investigation demonstrated that *NEDD9*, the only gene within the minimal common region that exhibited amplification-associated overexpression, was a *bona fide* melanoma metastasis oncogene.[145] Through the analysis of human and mouse models of B-cell lymphoma, Chang and colleagues demonstrated that *c-MYC* regulates a much broader set of miRNAs than previously anticipated. Notably, MYC over-expression promoted a widespread repression of miRNA expression, primarily through direct binding to miRNA promoters.[146] An important advantage of the simultaneous study of human and mouse tumors is that putative candidate genes can be functionally validated *in vivo.*

The identification of tumor-suppressor genes in cancer by classical genetics methods has been difficult and slow. In one report, integration of genomic and gene-expression microarray data was applied to localize suppressor genes. Within 20 homozygous deletion areas detected in 48 human B-cell lymphoma cell lines, a number of novel candidate genes were pinpointed.[100] Notably, some of these genes were shown to be inactivated in lymphoma biopsies by various genetic and epigenetic mechanisms that substantially varied among the different lymphoma subgroups. Thus, the P53-inducible *PIG7/LITAF* was silenced by homozygous deletion in primary mediastinal B-cell lymphoma and by promoter hypermethylation in germinal center lymphoma, whereas the proapoptotic *BIM* gene presented homozygous deletion in mantle cell lymphoma and promoter hypermethylation in Burkitt lymphoma.[100] A different study evaluated the candidate target genes in chromosome 8p21.3 deletions delineated through high-resolution array-CGH of B-cell lymphomas. In previous report, the presence of deletions of 8p in mantle cell lymphoma was associated with blood dissemination.[83,147] By comparing gene expression profiles of tumors with and without 8p deletion, authors found that only two genes within the 8p21.3 deletion, those encoding for the TRAIL receptors R1 and R2, showed significant downregulation in deleted tumors.[148] However, a recent report discovered that deletion of *BIN3*, another gene included within the 8p21.3 commonly deleted region, generated B-cell lymphoma in aging mice.[149] Loss of BIN3, which is a BAR adapter protein, did not affect normal cell proliferation but rather increased the motility of transformed cells. It is tempting to speculate that the loss of BIN3 may enhance B-cell lymphocyte migration, leading to a disseminated disease in patients with mantle cell lymphoma. A similar integrative microarray analysis revealed down-regulation of the gene encoding P53-binding protein 1 (*53BP1*) in DLBCL with heterozygous deletion of chromosome 15q15, being this deletion more common in the BCR-DLBCL group.[150] Although a reduced gene and protein dosage (haploinsufficiency) caused by the single-copy loss is suggested as the tumoral pathogenetic mechanism in these reports, further investigations are needed to validate this attractive hypothesis.

A different strategy combined nonsense-mediated RNA decay microarrays and array-CG for the genome-wide identification of genes with biallelic inactivation involving nonsense mutations and loss of the wild-type allele. This approach enabled the authors to identify previously unknown inactivating mutations in the receptor tyrosine kinase gene *EPHB2*, which were shown to be functionally important in the progression and metastasis of prostate cancer.[151] Zardo and colleagues used an alternative approach that integrated array-CGH and restriction landmark genomic scanning for global analysis of aberrant methylation of CpG islands in a series of human glioblastomas.[152] Results showed that most aberrant methylation events are focal and independent of genomic deletions, but a small subset of genes were affected by convergent methylation and deletion, including genes that exhibit tumor-suppressor activity such as such as *SOCS1* and *COE3*. In a different study, Stransky and colleagues used a combination of transcriptome correlation map analysis and array-CGH to evaluate at a large scale epigenetic suppression of gene expression of whole genomic regions. Authors demonstrated such regional copy number-independent deregulation of transcription by long-range epigenetic silencing in a series of bladder carcinomas.[153] In another study, authors determined the expression profiling of microRNAs in T24 cells, revealing that 17 out of 313 miRNAs were upregulated after DNA demethylation and histone deacetylase inhibition treatment. One of these, miR-127, was shown to repress *BCL6* oncogene, suggesting a role in the pathogenesis of this disease.[154]

Multiple myeloma is one of the tumors where integrative oncogenomic approaches have been more successfully applied. Shaughnessy and colleagues performed microarray analysis on myeloma cells from 532 patients. Seventy genes, 30% of them mapping to chromosome 1, were linked to reduced length of survival. Importantly, most up-regulated genes mapped to chromosome 1q (frequently amplified in myeloma), and down-regulated genes mapped to chromosome 1p (frequently deleted in myeloma). These data suggest that altered transcriptional regulation of chromosome 1 genes contribute to multiple myeloma pathogenesis and can be used to identify high-risk disease.[60] In a different study, high-resolution array-CGH data and expression profiles were determined in a collection of myeloma cell lines and patient biopsies.

Unsupervised classification defined distinct genomic subtypes. Genomic and expression data integration generated a refined list of myeloma gene candidates, thereby providing a molecular framework for dissection of disease pathogenesis.[155] More recently, two different groups investigated possible genetic lesions responsible for the constitutive NF-kB activation observed in multiple myeloma by integrating array-CGH and gene expression profiling data. Keats and colleagues found mutations in ten genes causing the inactivation of *TRAF2*, *TRAF3*, *CYLD*, *cIAP1*/*cIAP2* and activation of *NFKB1*, *NFKB2*, *CD40*, *LTBR*, *TACI*, and *NIK* that result primarily in constitutive activation of the noncanonical NF-kB pathway, with the single most common abnormality being inactivation of TRAF3.[156] Annunziata and colleagues compared the genetic profiles of multiple myeloma cell lines that were resistant or sensitive to an inhibitor of IkappaB kinase beta (IKKbeta) targeting the NF-kB pathway. Sensitive cell lines with NF-kB activation showed frequent genetic or epigenetic alteration of *NIK*, *TRAF3*, *CYLD*, *cAPI1*/*cAPI2*, *CD40*, *NFKB1*, or *NFKB2* genes.[157] These two complementary reports uncovered frequent genetic lesions of genes in the NF-kB pathway, suggesting that NF-kB inhibitors hold promise for the treatment of this disease.

## 5. FUTURE INVESTIGATIONS: INTEGRATIVE COMPUTATIONAL ANALYSIS OF NOVEL HIGH-THROUGHPUT GENETIC TECHNOLOGIES IN CANCER BIOLOGY

A miriad of new high-throughput technologies are being used in cancer research, including exon arrays to analyze alternative splicing, tiling arrays for high-resolution investigation of DNA and histone methylation patterns, on-chip chromatin immunoprecipitation to discover DNA-protein interactions and protein microarrays to measure global protein expression portraits. Consequently, next comparative oncogenomic and proteomic assays will attempt to viualize these complex molecular interactions in the context of highly connected and regulated cellular networks. While we assist to these fantastic advances, our last challenge is to use this comprehensive biological knowledge to accelerate the transition from current empirical therapies to tailored medicine.

## 6. MATERIALS

### TOTAL RNA PREPARATION FOR MICROARRAY ANALYSIS

This protocol is suitable for total RNA sample preparation for microarray analysis from cell lines or fresh frozen tissues. RNA obtained this way is very clean and salt free.

### MATERIALS

#### Reagents

- TRIzol® Reagent, Invitrogen Life Technologies, P/N 15596-018
- RNeasy® Mini Kit, QIAGEN, P/N 74104

  Contains:

  Rneasy® Mini Spin columns

  RLT buffer

  RPE buffer

  RNase-free water
- Absolute ethanol (store ethanol at room temperature)
- 80% ethanol (store ethanol at room temperature)

#### Supplies

- IKA® T-10 Basic Homogenizer (for fresh frozen tissue)
- Nanodrop ND-1000 Spectrophotometer
- 2100 Bioanalyzer and Agilent, RNA 6000 Nano LabChip® kit, P/N 5067-1511

### PROCEDURES

1a. For fresh frozen tissue samples: The amount of tissue required is variable depending on the kind of tissue and varies from 10 to 100 mg to get 10-300 µg of total RNA. Be careful not to let tissue thaw before homogenization. Homogenize tissue directly in TRIzol® reagent using an electric homogenizer by means of a small gauge generator (5 mm). Recommended volume of TRIzol® is 1 mL for each 50-100 mg of tissue. Homogenize each sample tube at least 3 times for at least 1 minute each time. Keep the samples on ice in between each

round of homogenization because overheat of samples can cause RNA degradation.

1b. For cell lines: Pellet cells by centrifugation and completely remove culture medium. Do not wash cells at this moment, proceed directly to lyse cells with the appropriate amount of TRIzol® reagent (recommended by manufacturer 1 mL/5-10×$10^6$ cells) by pipetting.

2. Let the samples stand for 5 minutes at room temperature.

3. Pass the sample twice through a 25G needle in order to reduce viscosity of the sample.

4. Add 200 µL of chloroform per mL of TRIzol® used and shake the sample for 15 s vigorously by hand. Incubate for 1 minute and shake again for 15 s.

5. Centrifuge the sample at 12,000 x *g* for 15 minutes at 2-8 ºC.

6. Following centrifugation the mixture separates into two phases, the colourless upper one is the aqueous phase where there is the RNA. The other one is the pink-phase (phenol-chloroform) which contains DNA and proteins. Take 200 µL from the top layer to continue and add them to 700 µL of QIAGEN RLT buffer in a new RNase-free tube. (Do not add 2-mercaptoethanol to RLT buffer because it may increase background in the array).

7. Add 500 µL of absolute ethanol to the sample (200 µL + 700 µL RLT). Mix well by vortex.

8. Apply the mixture to a QIAGEN Mini or MicroElute spin column and spin 15 s at 8,000 x *g*. Discard the flow through and repeat the procedure until all the sample has been loaded onto the column.

9. Replace the collector tube for a new one and wash the column by adding 500 µL of the RPE buffer. Centrifuge 15 s at 8,000 x *g* and discard flow through.

10. Add 700 µL of 80% ethanol and spin at 8,000 x *g* 15 s. Repeat this step again to efficiently remove all guanidine salts.

11. Transfer the column to a new collector tube and spin for 5 min at top speed with tubes cap off to ensure removal of ethanol.

12. To elute RNA, transfer the column to a new 1.5 mL RNase-free microfuge tube. Elute with 20 µL or 14 µL of RNase-free water for Mini or MicroElute Spin column, respectively.

**QUALITY CONTROL OF RNA**

To qualify RNA for microarray applications it is important to measure its concentration, 260/280 ratio, 260/230 ratio and RNA integrity. We use Nanodrop to asses that the concentration is at least 250 ng/µL, the 260/280 ratio is between 1.9-2.1; and the 260/230 ratio is greater than 1.5 (this determines the presence of salts that could inhibit labelling reactions). Integrity of RNA can be measured by studying integrity of rRNA on gel. Affymetrix recommends the use of the capillary electrophoresis Bioanalyzer 2100 system from Agilent. This software calculates the RIN (RNA Integrity Number), which in our experience should be greater than 8.0 to guarantee that the sample will work properly on the array.

## DNA PREPARATION FOR MICROARRAY ANALYSIS

This protocol is based on the procedure established by QIAGEN using their DNeasy® Blood & Tissue kit.

### MATERIALS

#### Reagents

- DNeasy® Blood & Tissue kit, QIAGEN, P/N 69504

  Contains:

  Dneasy® Mini Spin columns

  ATL buffer

  Proteinase K

  AL buffer

  AW1 buffer

  AW2 buffer
- Absolute ethanol
- Reduced EDTA TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0)

#### Supplies

- Nanodrop ND-1000 Spectrophotometer

### PROCEDURE

<u>For tissue samples</u>:

1a. The amount of tissue needed is variable but 25 mg tissue (up to 10 mg spleen) maybe suitable for this application. Cut the tissue into small pieces, and place it in a 1.5 mL microcentrifuge tube. Add 180 µL Buffer ATL.

2a. Add 20 µL proteinase K. Mix thoroughly by vortexing, and incubate at 55 °C until the tissue is completely lysed (it can be lysed overnight). During incubation it is recommended to vortex occasionally to disperse the sample.

3a. Add 200 µl Buffer AL to the sample, and mix thoroughly by vortexing. Then add 200 µL ethanol (96–100%), and mix again by vortexing. It is essential that the sample, Buffer AL, and ethanol are mixed immediately and thoroughly to yield a homogeneous solution.

<u>For cell lines</u>:

1b. Start from around $5x10^6$ cells, pellet them and wash twice with 1X PBS. Resuspend pellet in 200 µL 1X PBS.

2b. Add 20 µL proteinase K and 200 µL of buffer AL mix thoroughly by vortexing and place at 70 °C for 10 minutes.

3b. Then add 200 µL ethanol (96–100%), and mix again thoroughly by vortexing.

4. Pipet the sample (including any precipitate) into the DNeasy® Mini spin column placed in a 2 mL collection tube. Centrifuge at 6,000 x *g* for 1 min. Discard flow-through and collection tube.

5. Place the DNeasy® Mini spin column in a new 2 mL collection tube, add 500 µL Buffer AW1, and centrifuge for 1 min at 6,000 x *g*. Discard flow-through and collection tube.

6. Place the DNeasy® Mini spin column in a new 2 mL collection tube, add 500 µL Buffer AW2, and centrifuge for 3 min at 20,000 x *g* to dry the DNeasy® membrane. Discard flow-through and collection tube.

7. Place the DNeasy® Mini spin column in a clean 1.5 mL or 2 mL microcentrifuge tube, and pipet 200 µL Buffer AE directly onto the DNeasy® membrane. Incubate at room temperature for 1 min, and then centrifuge for 1 min at 6,000 x *g* to elute. If SNP arrays from Affymetrix are to be performed, then use a buffer with lower EDTA concentration to elute the

sample (10 mM Tris-HCl; 0.1 mM EDTA pH 8.0) as EDTA concentration adversely affects following reactions.

**QUALITY CONTROL OF DNA**

Principal parameters to control DNA quality are concentration (for 500K SNP array from Affymetrix it should be at least 50 ng/µL), 260/280 ratio around 1.9 if pure DNA and 260/230 ratio greater than 1.5 in salt-free samples. To determine DNA integrity, we perform gel electrophoresis on a 1-2% agarose 1X TBE gel. High quality genomic DNA will give a band of 10-20 Kb on the gel.

## 7. METHODS FOR RNA AND DNA MICROARRAYS

## Oligonucleotide gene expression microarrays

Introduction

We use the One-Cycle Eukaryotic Target Labelling Assay from Affymetrix. It is possible to start with total RNA (1 µg to 15 µg) or mRNA (0.2 µg to 2 µg). We normally begin with 2 µg of total RNA. It is fundamental to start with the same amount of RNA for all samples to be compared. This RNA is first reverse transcribed using a T7-Oligo(dT) Promoter Primer. Second-strand synthesis reaction is mediated by RNase H. Double-stranded cDNA obtained is then purified and used as a template in the following *in vitro* transcription (IVT) reaction. The IVT reaction is performed in the presence of T7 RNA Polymerase and a biotinylated nucleotide analog/ribonucleotide mix for complementary RNA (cRNA) amplification and biotin labelling. This biotinylated cRNA targets are then cleaned up, fragmented, and hybridized to GeneChip expression arrays.

### Materials

**Reagents**

- **Labelling**
  - One-Cycle Target Labelling and Control Reagents, Affymetrix, P/N 900493
  - Contains:
  - 1 IVT Labelling Kit (Affymetrix, P/N 900449)
  - 1 One-Cycle cDNA Synthesis Kit (Affymetrix, P/N 900431)
  - 1 Sample Cleanup Module (Affymetrix, P/N 900371)
  - 1 Poly-A RNA Control Kit (Affymetrix, P/N 900433)
  - 1 Hybridization Control Kit (Affymetrix, P/N 900454)
  - Absolute ethanol
  - 80% ethanol

**Hybridization, Stain and Wash**

- GeneChip Hybridization, Wash and Stain kit, Affymetrix, P/N 900720 (30 reactions)
- Contains:
- Hybridization module (Pre-Hybridization mix, 2X Hybridization mix, DMSO, Nuclease-free water)
- Stain Module (Stain Cocktail 1, Stain Cocktail 2, Array Holding Buffer)
- Wash Buffers (Wash Buffer A, Wash Buffer B)
- GeneChip Eukaryotic Hybridization Control Kit, Affymetrix, P/N 900454 (30 reactions) or P/N 900457 (150 reactions), contains Control cRNA and Control Oligo B2

**Supplies**

- Nanodrop ND-1000 Spectrophotometer
- 2100 Bioanalyzer and Agilent, RNA 6000 Nano LabChip® kit, P/N 5067-1511
- Hybridization Oven 640, Affymetrix, P/N 800138 (110V) or 800139 (220V)
- Heatblock
- Fluidics Station 450: Affymetrix, P/N 00-0079
- GeneChip® Scanner 3000: Affymetrix, P/N 00-00212

## Procedures

### Preparation of Poly-A RNA Controls for One-Cycle cDNA Synthesis (Spike-in Controls)

Relative amount of Poly-A RNA Controls added to the sample RNA will be constant, therefore it is dependent on the initial µg of sample. For 2 µg of RNA, 2 µL of a 1:50,000 dilution of Poly-A RNA Controls are used.

### First-Strand cDNA Synthesis

1. Mix RNA sample, diluted poly-A RNA controls, and T7-Oligo(dT) Primer. Incubate the reaction for 10 minutes at 70 ℃. Then cool the sample at 4 ℃ for at least 2 minutes.
2. In a separate tube, assemble the First-Strand Master Mix: 4.0 µL 5X 1st Strand Reaction Mix; 2.0 µL 0.1 M DTT; 1 µL 10 mM dNTP (per sample).

3. Transfer 7 µL of First-Strand Master Mix to each RNA/T7-Oligo(dT) Primer mix for a final volume of 19 µL. Mix by flicking the tube a few times. Immediately place the tubes at 42 °C and incubate for 2 minutes at 42 °C.
4. Add 1 µL of SuperScript II to each RNA sample for a final volume of 20 µL.
5. Incubate for 1 hour at 42 °C; then cool the sample for at least 2 minutes at 4 °C.

**Second-Strand cDNA Synthesis**

1. Prepare Second-Strand Master Mix: 91 µL RNase-free Water; 30 µL 5X 2nd Strand Reaction Mix; 3 µL 10 mM dNTP; 1 µL *E. coli* DNA ligase; 4 µL *E. coli* DNA Polymerase I; 1 µL RNase H (per sample).
2. Add 130 µL of Second-Strand Master Mix to each first-strand synthesis sample from First-Strand cDNA Synthesis for a total volume of 150 µL. Then incubate for 2 h at 16 °C.
3. Add 2 µL of T4 DNA Polymerase to each sample and incubate for 5 additional minutes at 16 °C.
4. Then add 10 µL 0.5 M EDTA and proceed to Cleanup of Double-Stranded cDNA. Do not leave the reactions at 4 °C for long periods of time.

**Cleanup of Double-Stranded cDNA**

1. Add 600 µL of cDNA Binding Buffer to the double-stranded cDNA synthesis preparation and mix by vortexing for 3 seconds. Colour of the mixture should be yellow. If not add 10 µL of 3 M sodium acetate pH 5.0, and mix.
2. Apply 500 µL of the sample to the cDNA Cleanup Spin Column sitting in a 2 mL Collection Tube, and centrifuge for 1 minute at ≥ 8,000 x *g*. Discard flow-through. Repeat reload of the spin column with the remaining mixture and centrifuge as above. Discard flow-through and Collection Tube.
3. Transfer spin column into a new 2 mL Collection Tube. Wash spin column with 750 µL of the cDNA Wash Buffer. Centrifuge for 1 min at ≥ 8,000 x *g*. Discard flow-through.
4. Open the cap of the spin column and centrifuge for 5 min at maximum speed to completely eliminate ethanol. Discard flow-through and Collection Tube.

5. Transfer spin column into a 1.5 mL Collection Tube, and pipet 14 µL of cDNA Elution Buffer directly onto the spin column membrane. Incubate for 1 minute at room temperature and centrifuge 1 min at maximum speed (≤25,000 x *g*) to elute.

**Synthesis of Biotin-Labeled cRNA**

1. Transfer the needed amount of template cDNA (if 2 µg were used as starting material, use 12 µL of purified cDNA) to RNase-free microfuge tubes and add the following reaction components in the order indicated: 8 µL RNase-free Water; 4 µL 10X IVT Labelling Buffer; 12 µL IVT Labelling NTP Mix; 4 µL IVT Labelling Enzyme Mix. It is important not to assemble the reaction on ice, because spermidine in the 10X IVT Labeling Buffer can lead to precipitation of the template cDNA.
2. Incubate at 37 ℃ for 16 h in a thermal cycler.

**Cleanup and Quantification of Biotin-Labeled cRNA**

1. Add 60 µL of RNase-free Water to the IVT reaction and mix by vortexing for 3 s.
2. Add 350 µL IVT cRNA Binding Buffer to the sample and mix by vortexing for 3 s.
3. Add 250 µL ethanol (96-100%) to the lysate, and mix well by pipetting. Do not centrifuge at this step.
4. Apply sample (700 µL) to the IVT cRNA Cleanup Spin Column sitting in a 2 mL Collection Tube. Centrifuge for 15 s at ≥ 8,000 x *g*. Discard flow-through and Collection Tube.
5. Transfer the spin column into a new 2 mL Collection Tube. Pipet 500 µL IVT cRNA Wash Buffer onto the spin column. Centrifuge for 15 seconds at ≥ 8,000 x *g* to wash. Discard flow-through.
6. Pipet 500 µL 80% (v/v) ethanol onto the spin column and centrifuge for 15 seconds at ≥ 8,000 x *g*. Discard flow-through.
7. Centrifuge for 5 minutes with caps off at maximum speed to allow complete drying of the membrane. Discard flow-through and Collection Tube.

8. Transfer spin column into a new 1.5 mL Collection Tube, and pipet 21 µL of RNase-free Water directly onto the spin column membrane. Centrifuge 1 min at maximum speed (≤ 25,000 x *g*) to elute.

For subsequent quantification of the purified cRNA, we dilute the eluate 1:5 or 1:4 fold in RNase-free water. We use Nanodrop to determine concentration of the cRNA obtained and Bioanalyzer to study sizes of the labelled products (which should have an average size of 1580 nucleotides).

If using total RNA as starting material, it is necessary to calculate an adjusted cRNA yield to reflect carryover of unlabeled total RNA. Using an estimate of 100% carryover, use the formula below to determine adjusted cRNA yield:

$$adjusted\ cRNA\ yield = RNAm - (total\ RNAi)\ (y)$$

*RNAm* = amount of cRNA measured after IVT (µg)

*total RNAi* = starting amount of total RNA (µg)

*y* = fraction of cDNA reaction used in IVTSample Cleanup Module

**Fragmenting the cRNA for Target Preparation**

1. Fragmentation of cRNA is a critical step of the protocol. When using a 49 microarray format, we will fragment 20 µg (with a volume ranging from 1 to 21 µL). Final volume of fragmentation reaction is 40 µL where 8 µL correspond to 5X Fragmentation Buffer.
2. Incubate reaction at 94 ℃ for 35 minutes. Put on ice following the incubation.

Save an aliquot for analysis on the Bioanalyzer. This standard fragmentation procedure should produce a distribution of RNA fragment sizes from approximately 35 to 200 bases.

Undiluted, fragmented sample cRNA is ready to perform the hybridization. If you are not going to proceed with labelling at the moment store the sample at -20 ℃ (or -70 ℃ for longer-term storage).

**Hybridization**

1. Mix the following for each target, scaling up volumes for hybridization to multiple probe arrays.

15 µg Fragmented cRNA (final concentration 0.05 µg/µL)

5 µL Control Oligonucleotide B2 3 nM (final concentration 50 pM)

15 µL 20X Eukaryotic Hybridization Controls (*bioB*, *bioC*, *bioD*, *cre*) (final concentration 1.5 pM)

150 µL 2X Hybridization Buffer (final concentration 1X)

30 µL DMSO (final concentration 10%)

Nuclease-free water up to 300 µL

2. Equilibrate probe array to room temperature immediately before use.
3. Heat the hybridization cocktail to 99 ℃ for 5 minutes in a heat block.
4. Meanwhile, wet the array by filling it through one of the septa with appropriate volume of 1X Pre-Hybridization Buffer using a micropipettor and appropriate tips. Incubate the probe array at 45 ℃ for 10 minutes with rotation.
5. Transfer the hybridization cocktail that has been heated at 99 ℃, in step 3, to a 45 ℃ heat block for 5 minutes.
6. Spin hybridization cocktail(s) at maximum speed in a microcentrifuge for 5 minutes to remove any insoluble material from the hybridization mixture.
7. Remove the buffer solution from the probe array cartridge and fill with 200 µL (for the 49-format) of the clarified hybridization cocktail, avoiding any insoluble matter at the bottom of the tube.
8. Place probe array into the Hybridization Oven, set to 45 ℃. Avoid stress to the motor; load probe arrays in a balanced configuration around the axis. Rotate at 60 r.p.m. Hybridize for 16 hours.

**Staining, washing and scanning**

Staining and washing are performed using the Fluidics Station 450 (Affymetrix). At this point, the most important is to select the correct script for your chip. For example HUG-133 2.0 Plus uses protocol FS450_0001. The script contains the directions to stain and wash the microarray: number of cycles of wash or stain, temperature, buffer.

For HUG-133 2.0 Plus place Stain Cocktail 1 in sample holder 1, Stain Cocktail 2 in sample holder 2 and Array Holding Buffer in sample holder 3. In the final step, probe array is filled with array holding buffer; arrays can be stored for 3 hours at 4 ℃ at dark before scanning.

The scanner used is GeneChip® Scanner 3000. Complete image of scanned array is stored as a .DAT file (scanned image, full information), and then GCOS software generates the .CEL file which represents the first summarization step because image is summarized in median intensity/probe cell.

## CGH to BAC microarrays

### Introduction

The arrays for CGH consist of linker-adapter PCR representation of BAC clones printed on a substrate. Each clone contains at least one STS (Sequence Tagged Site) and is mapped to the human genome sequence. Clones containing unique sequences near telomeres and clones containing genes known to be significant in cancer and medical genetics are included. Hybridization to these arrays allows detection of single copy gains and losses compared to diploid cells even in presence of normal cell contamination.

### Materials

**Random primed labeling of genomic DNA for array CGH analysis**

- 2.5X Random Primers (BioPrime DNA labeling systems, Invitrogen 18094-011). Store a -20 ℃.
- Genomic DNA.
- Klenow fragment (40 U/ µL, BioPrime DNA labeling system, Invitrogen 18094-011). Store a -20 ℃.
- Cy3 and Cy5 labeled dCTP (1 mM, Amersham Pharmacia Biotech Inc. PA53021 and PA55021).
- 0.5 M EDTA, pH 8.0.
- 1 M Tris-HCl, pH 7.6.
- 10X dNTP mixture in sterile water: 3.7 mM dATP, dTTP and dGTP (Invitrogen, 10216-018, 10219-012 and 10218-014, respectively), 1.8 mM dCTP (Invitrogen 10217-016), 10 mM Tris-HCl pH 7.6 and 1 mM EDTA.
- Sephadex G-50 spin column (Amersham Pharmacia Biotech Inc. 27-5330-01).

**Hybridization of fluorescently labeled genomic DNA for array CGH analysis**

- Human cot-1 DNA (1 mg/µL, Invitrogen)
- 20% SDS in sterile $H_2O$ (heat at 68 ℃ to dissolve).
- 100% Ethanol. Store ay -20 ℃.
- 3.0 M Sodium acetate, pH 5.2.
- Dextran sulfate sodium salt (500,000 MW).
- Formamide (re-destilled, ultra pure, Invitrogen). Store a -20 ℃
- 20X SSC (3.0 M NaCl, 0.3 M sodium citrate, pH 7.0).
- Mastermix mixture: dissolve 1 g dextran sulfate in 5 mL of formamide, 1 mL of 20X SSC and 1 mL $dH_2O$. Adjust to pH 7.0 with approximately 2 drops of HCl.
- PN buffer: 0.1 M sodium phosphate, 0.1% Nonidet P40, pH 8.0.
- UV Stratalinker 2400 (Stratagene) capable of producing 130,000 x 100 µJoules UV.
- Very slow rocking table (-1 rpm) inside a 37 ℃ incubator.
- Rubber cement (Ross, American Glue Corporation).
- Silicon gasket (Press-to-seal, 2 mm thick, #62-6508-24, PGC Scientific).
- 100% Glycerol
- 10X PBS
- Stereomicroscope.
- Binder clips, medium size.

**Supplies**

- UV Stratalinker 2400 (Stratagene)
- 1M Pixel CCD Imager (custom made; Dan Pincel, UCSF) or the 2-color scanner arrayWoRxe Biochip Reader (AppliedPrecision, Issaquah, Washington, USA), a white-light CCD-based system that provides highest quality images along with more accurate and repeatable microarray results.

**<u>Procedures</u>**

**Random primed labeling of genomic DNA for array CGH analysis**

A typical random primed labeling procedure is described. The random primed labeling is carried out in a 25 µL reaction volume containing: 600 ng genomic DNA, 1X random primers, 40 U Klenow DNA polymerase, Cy3 and Cy5 labeled dCTP and 1X dNTP mixture.

1. Mix 6000 ng genomic DNA with 10 µL of 2.5X random primer solution and make up the volume to 21 µL with sterile $H_2O$.
2. Denature the DNA by heating the mixture at 99 °C in a PCR machine for 10 min. Briefly centrifuge and place on ice.
3. Add: 2.5 µL of the 10X dNTP mixture, 1 µL of 1 mM Cy3 and Cy5 labeled dCTP and 0.6 µL Klenow DNA polymerase. Incubate at 37 °C for 12-20 h.
4. Remove unincorporated nucleotides from the DNA.
   Place a Sephadex G-50 column in a 1.5 mL tube and pre-spin the column at 760 x *g* for 1 min. Discard the supernatant. Tap the end of the tube to a paper towel to remove the remaining supernatant from the neck of the tube. Place the column in a clean 1.5 mL tube, apply the sample onto the column and spin at 760 x *g* for 2 min.

**Hybridization of fluorescently labeled genomic DNA for array CGH analysis**

1. Preparation of the array for the hybridization:
   a. Expose a printed array to 260,000 µJ (2,600 x 100 µJ) of UV by using a Stratalinker. Place the slide in the Stratalinker, array facing up. Overcrosslinking the slide might result in a decrease in fluorescent hybridization signal.
   b. Fill a 10 mL syringe with rubber cement and fit a 200 µL pipet tip on the syringe outlet. You may have to cut 1-2 mm off the wide of the pipet tip for it to fit well. Apply a rubber cement ring around each array on the slide using a stereomicroscope to observe the area of the array. Air-dry and apply a second thick layer of rubber cement on top of the first layer Air-dry the rubber cement.
2. Preparation of samples for hybridization:

a. Combine 25 µL labeled test genomic DNA, 25 µL labeled reference genomic DNA and 40-50 µg of human Cot-1 DNA. Precipitate the DNA sample mixture by adding 2.5 volumes of ice-cold 100% ethanol and 0.1 volume of 3 M sodium acetate pH 5.2. Vortex the solution briefly and collect the precipitate by centrifugation at 14,000 x *g* for 45 min at 4 ℃.

b. Carefully aspirate and discard the supernatant. Wipe the excess liquid from the tube and air-dry the pellet for approximately 5-10 min. Dissolve the pellet in 7 µL $dH_2O$, 14 µL 20% SDS, and 49 µL Master mix mixture. Incubate 1 h at room temperature to completely resuspend.

3. Denature the DNA sample at 73 ℃ for 13 min and then incubate at 37 ℃ for 1-2 h to allow the Cot-1 DNA to anneal to repetitive sequences.
4. Place array on a heat block set at 37 ℃ for 5 min to warm the array.
5. Apply the sample (step 3) onto the array. Keep the sample at 37 ℃ until just before application to the array to reduce non-specific binding of the probe to the array surface. Place a silicon gasket around the edge of the slide and lay a clean glass slide on top, aligning the edges with the gasket. Clamp the assembly together using binder clips. Incubate the array for 48-68 h at 37 ℃ on a slowly rocking table (--1 r.p.m.).
6. Disassembly the array assembly and rinse the hybridization solution from the slide under a stream of PN buffer. It is preferred to leave the rubber cement on the array at this time, as it will not affect the rising steps that follow.
7. Wash the slides once on 50% formamide, 2X SSC, pH 7.0 for 15 min at 45 ℃, followed by a 15 min wash in PN buffer at room temperature. The washes can conveniently be done in slide staining jars (coupling jars) placed on water baths.
8. At the bench, carefully remove the rubber cement with forceps, while keeping the array moist with PN buffer.
9. Mount the slide in a DAPI solution to stain the array spots (90% glycerol, 10% PBS, 1 µM DAPI).

**Microarray image capture with CCD Imager and microarray image quantification**

To capture the microarray image with CCD Imager for the image quantification we use the software "UCSF SPOT" available in www.janlab.org/downloads.htmL. This software allows the obtaining of numerical values, expressed in $\log_2$ ratio, for the ratios comprised between the sample to be analyzed and the control sample. The numerical data are processed and saved in an Excel table. Using the software "SPROC", the data are normalized from the .spot files generating the final $\log_2$ ratio file data with the standard deviation (media of each three spots). At the same time, the program arranges the BACs by its genomic position and chromosome location (http//genome.vse.ucsc.edu)

## High resolution SNP-CGH microarrays

### Introduction

The purpose of the Affymetrix GeneChip Mapping 500K Assay is to detect Single Nucleotide Polymorphisms (SNPs) greater than 500,000 in samples of genomic DNA. The Mapping 500K Set is comprised of two arrays and two assay kit. The protocol starts with 250 ng of genomic DNA per array and will generate SNP genotype calls for approximately 250,000 SNPs for each array of the two array set. The assay utilizes a strategy that reduces the complexity of the human genomic DNA up to 10 fold by first digesting the genomic DNA with the *Nsp*I or *Sty*I restriction enzyme and then ligating sequences onto the DNA fragments. The complexity is further reduced by a PCR procedure optimized for fragments of a specified size range. After these steps the PCR products are fragmented, end-labeled, and hybridized to a Gene Chip array.

### Materials

- Reduced EDTA TE Buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0): TEKnova, P/N T0223.

- 250 ng Genomic DNA per array working stock: 50 ng/µL.

- *Sty*I (10,000 U/mL): New England Biolabs (NEB), P/N R0500S
- *Nsp*I (10,000 U/mL): New England Biolabs (NEB), P/N R0602L
- AccuGENE® Water, Molecular Biology-Grade, Cambrex, P/N 51200
- T4 DNA Ligase: New England Biolabs (NEB), P/N M0202L
- Adaptor Nsp (50 µM): Affymetrix, P/N 900596 for 30 Rxns and P/N 900697 for 100 Rxns.
- Adaptor Sty (50 µM): Affymetrix, P/N 900597 for 30 Rxns and P/N 900698 for 100 Rxns.
- G-C Melt (5 M): Clontech, P/N 639238
- dNTP (2.5 mM): Takara, P/N 4030, or Fischer Scientific, P/N TAK 4030
- PCR Primer 002 (100 µM): Affymetrix, P/N 900595 for 30 Rxns and P/N 900702 for 100 Rxns.
- Clontech TITANIUM® Taq Polymerase (50x): Clontech, P/N 63920.
- All purpose Hi-Lo DNA Marker: Bionexus, Inc., P/N BN2050
- DNA amplification Clean-up kit, to be used with Affymetrix DNA products: Clontech, P/N 636974 (1 plate). The kit contains RB Buffer.
- Fragmentation Reagent (DnaseI): Affymetrix, P/N 900131
- 10x Fragmentation buffer: Affymetrix, P/N 900422 for 30 Rxns
- 4% TBE Gel: BMA Reliant precast (4% NuSieve 3:1 Plus Agarose): Cambrex, P/N 54929
- GeneChip® DNA Labeling Reagent (30 mM): Affymetrix, P/N 900778 for 30 Rxns and P/N 900699 for 100 Rxns.
- Terminal Deoxynucleotidyl Transferase (30 U/µL): Affymetrix, P/N 900508 for 30 Rxns and P/N 900703 for 100 Rxns.
- 5X Terminal Deoxynucleotidyl Transferase Buffer: Affymetrix, P/N 900425 for 30 Rxns and P/N 900696 for 100 Rxns.
- 5 M TMACL (Tetramethyl Ammonium Chloride): Sigma, P/N T3411
- MES Hydrate Sigma Ultra: Sigma, P/N M5287
- MES Sodium SALT: Sigma, P/N M5057
- Denhardt's Solution: Sigma, P/N D2532
- HSDNA (Herring Sperm DNA): Promega, P/N D1815

- Human Cot-1 DNA®: Invitrogen, P/N 15279-011
- Oligo control reagent, 0100 (OCR, 0100): Affymetrix. P/N 900541 for 30 Rxns and 900701 for 100 Rxns.
- GeneChip 250K Array (one per sample).

**Supplies**

- GeneAmp® PCR System 9700 Thermocycler by Applied Biosystems
- GeneChip Hybridization oven 640.
- Manifold-QIAvac multiwell unit: QIAGEN, P/N 9014579
- Biomek® Seal and Sample aluminum foil lids: Beckman, P/N 538619
- Jitterbug® 115 VAC: Boekel Scientific, P/N 130000
- QIAGEN® Vacuum regulator: QIAGEN, P/N 19530

**Procedures**

**STEP 1. Genomic DNA Preparation**

To minimize contamination of the simples, the use of two separate rooms to perform the assay is recomended: one is the pre-PCR clean room (or area for the DNA template and free of PCR products), and the other is the PCR-Staging room or main lab, where the rest of steps are performed.

1. Thoroughly mix the genomic DNA by vortexing at high speed for 3 sec.
2. Determine the concentration of each genomic DNA sample.
3. Based on OD measurements, dilute each sample to 50 ng/µL using reduced EDTA TE buffer.

**STEP 2. Restriction enzyme digestion**

Before proceeding:

- Program the thermal cycler in advance. Switch on the thermal cycler 10 minutes before reactions are ready so that the lid is heated.
- Reference Genomic DNA 103 is supplied in both the Sty and NSP GeneChip Mapping 250K Assay Kits. This DNA can be used as a positive control.

1. Depending on the restriction enzyme used, prepare the following Digestion Master Mix ON ICE (for multiple samples, make a 5% excess): for the *Nsp*I digestion, 9.75 µL $H_2O$; 2 µL 10X NE Buffer 2; 2 µL 10X BSA (1 mg/mL); 1 µl *Nsp*I (10 U/µl). For the *Sty*I digestion, 9.75 µL $H_2O$; 2 µL 10X NE Buffer 3; 2 µL 10X BSA (1 mg/mL); 1 µL *Sty*I (10 U/µl).
Note: The BSA is supplied as 100x (10 mg/mL), and needs to be diluted 1:10 with molecular grade water before use.
2. Add 5 µL of genomic DNA diluted to each tube. Total amount of genomic DNA is 250 ng for each restriction enzyme
3. Aliquot 14.75 µL of the digestion master mix to each tube containing DNA. Mix gently and spin at 400 x *g*.
4. Place the tubes in the thermal cycler and run the *500K Digest* program: 37 °C, 120 min; 65 °C, 20 min; hold at 4 °C.
Store the sample at -20 °C if not proceeding to the next step.

**STEP 3. Ligation**

Before proceeding:
- Program the thermal cycler in advance. Switch on the thermal cycler 10 min before reactions are ready so that the lid is heated.
- Ligase buffer contains ATP and should be thawed/held at 4 °C. Avoid multiple freeze-thaw cycles, according to vendor's intructions.

1. Depending on the restriction enyme used, in the pre-PCR area prepare the following Ligation Master Mix ON ICE (for multple samples, prepare a 5% excess): for *Nsp*I, 0.75 µL Adaptor Nsp 50 µM; 2.5 µL 10X T4 DNA Ligase Buffer; 2 µL T4 DNA Ligase (400 U/ µl). For *Sty*I: 0.75 µL Adaptor Sty 50 µM; 2.5 µL 10X T4 DNA Ligase Buffer; 2 µL T4 DNA Ligase (400 U/ µl). Total volume: 5.25 µL.
2. Aliquot 5.25 µL of the Ligation Master Mix into each digested DNA simple. Add the 19.75 µL of the digested DNA until a total volume of 25 µL. Mix gently and spin at 400 x *g* for 1 min at 4 °C

3. Place tubes in a thermal cycler and run the *500K Ligate* program: 16 °C, 180 min; 70 °C, 20 min; hold at 4 °C.
Store samples at -20 °C if not proceeding to the next step within 60 minutes.
4. Dilute each DNA ligation reaction by adding 75 µL of molecular biology grade water to the 25 µL (1/4 dilution).

**STEP 4: PCR**

Before proceeding:
- Program the thermal cycler in advance. Switch on the thermal cycler 10 minutes before reactions are ready so that the lid is heated.

1. Prepare the following PCR master Mix ON ICE (3 PCR reactions per sample) in the pre-PCR clean room for *Nsp*I or *Sty*I ligation reactions and vortex at medium speed for 2 seconds (for multiple samples, make a 5% excess): for one PCR: 39.5 µL $H_2O$; 10 µL 10X Clontech TITANIUM® Taq PCR Buffer; 20 µL 5 M G-C Melt; 14 µL 2.5 mM dNTPs; 4.5 µL 100 µM PCR Primer 002; 2 µL 50X Clontech TITANIUM® Taq Polymerase.
Note: 90 µg of PCR product is needed for fragmentation
2. Transfer 10 µL of each diluted ligated DNA to the corresponding 3 PCR tubes.
3. Add 90 µL PCR Master mix to obtain a total volume of 100 µL.
4. Mix gently and spin samples at 400 x *g* for 1 min.
5. Place in the thermal cycler in the main lab and run the *500K PCR* program (optimized for the GeneAmp® PCR System 9700 Thermocycler): 94 °C, 3 min; 30x (94 °C, 30 s; 60 °C, 45 s; 68 °C, 15 s); 68 °C, 7 min; hold at 4 °C.
6. Run 3 µL of each PCR product mixed with 3 µL 2X Gel Loading Dye on 2% TBE gel at 120 V for 1 h.
PCR products can be stored at -20 °C if not proceeding to the next step within 60 min.

**STEP 5: PCR purification and elution with Clontech Clean-up plate**

1. Connect a vacuum manifold to a suitable vacuum source able to maintain 600 mbar aprox.
2. Place a Clean-Up Plate on top of the manifold. Cover wells that are not needed with PCR plate cover. We recommend to cover the plate with the aluminum cover, and remove the portion of the cover corresponding to the probe wells.
3. Add 8 µL of 0.1 M EDTA (diluted from the 0.5 M EDTA in water) to each PCR reaction. Seal plate with plate cover, vortex at medium speed for 2 seconds, and spin at 400 x *g* for 1 min.
4. Cosolidate three PCR reactions for each sample into one well of the Clean-Up plate.
5. Apply a vacuum and maintain at 600 mbar until the wells are completely dry.
6. Wash the PCR products by adding 50 µL molecular biology grade water and dry the wells completely (~ 20 min). Repeat this step 2 additional times for a total of three water washes.
7. Switch off vacuum source and release the vacuum.
8. Carefully remove the Clean-Up plate from the vacuum manifold and inmediately:
a) blot the plate on a stack of clean absorbent paper to remove any liquid that might remain on the bottom of the plate.
b) dry the bottom of each well with an absorbent wipe.
9. Add 45 µL RB buffer to each well. Cover the plate with PCR plate cover film and seal tightly. Moderately shake the Cean-Up Plate on a Plate shaker, for 10 minutes at room temperature.
10. Recover the purified PCR product to clean tubes by pippeting the eluate out of each well and transferring it to the corresponding tube.

**STEP 6: Quantification of purified PCR products**

1. Add 2 µL of the purified PCR product to 198 µl molecular biology grade water and mix well.
2. Read the absorbance at 260 nm. Ensure that the reading is in the quantitative range of the instrument (generally 0.2 to 0.8 OD).

3. Apply the convention* that 1 absorbance unit at 260 nm equals 50 µg/mL for double-stranded PCR products.

* This convention assumes a path length of 1 cm. Consult your spectrophotometer handbook for further information.

4. For fragmentation:

a) Transfer 90 µg of each of the purified DNA samples to the corresponding wells of a new plate.

b) Bring the total volume of each well up to 45 µL by adding the appropiate volume of RB buffer.

c) Cover the plate with PCR plate cover film and seal tightly.

d) Vortex at medium speed for 2 s, and spin down at 400 x *g* for 1 min.

**STEP 7: Fragmentation**

Before proceeding:

- Pre-heat the thermal cycler to 37 ℃ before setting up the fragmentation reaction
- Prepare the fragmentation dilution inmediately prior to use
- Perform all the dilution, and mixing steps on ice

1. Pre-heat thermal cycler to 37 ℃

2. Add 5 µL 10x Fragmentation Buffer to each sample (45 µL) in the corresponding tube ON ICE, giving a total volume of 50 µL.

3. Examine the label of the GeneChip Fragmentation Reagent tube for U/µl definition, and calculate dilution.

Y = number of µl of stock Fragmentation Reagent

X = number of U of stock Fragmentation Reagent per µL (see label on the tube)

0.05 U/µL = final concentration of diluted Fragmentation Reagent

120 µL = final volume of diluted Fragmentation Reagent (enough for 20 Rxns)

$$Y = 0.05\ U/\mu L * 120\ \mu L / x\ U/\mu L$$

4. Dilute the stock of Fragmentation Reagent to 0.05 U/µL as follows:

a) Place the water, Fragmentation Buffer and Fragmentation Reagent on ice

b) Combine the reagents ON ICE in the order shown below.

c) Vortex at medium speed for 2 s.

See an example of dilution as follow: 105 µL $H_2O$; 12 µL 10X Fragmentation Buffer; 3 µL Fragmentation Reagent; giving a total volume of 120 µL

5. Divide the Fragmentation Reagent in the tubes required

6. Add 5 µL of diluted Fragmentation Reagent (0.05 U/µL) to the PCR samples tubes containing Fragmentation mix on ice. Pipet up and down several times to mix. The total volume for each sample is 50 µL.

7. Mix gently the tubes and spin briefly at 400 x *g* at 4 ºC

8. Place the samples in a pre-heated thermo cycler as quickly as possible, and run the *500K Fragment* program: 37 ºC, 35 min; 95 ºC, 15 min; hold at 4 ºC.

9. Spin the samples to collect at the bottom of the tube

10. Dilute 4 µL of fragmented PCR product with 4 µL gel loading dye and run on 4% TBE gel.

Proceed inmediately to the Labeling step, if the result matches the example below.

**STEP 8: Labeling**

Before proceeding:

- Program the thermal cycler in advance. Switch on the thermal cycler 10 minutes before reactions are ready so that the lid is heated.

1. Prepare Labeling Mix ON ICE and vortex at medium speed for 2 seconds (for multiple samples, make a 5% excess): 14 µL 5X TdT Buffer; 2 µL 30 mM GeneChip® DNA Labeling Reagent; 3.5 µL TdT (30 U/µL).

2. Aliquot 19.5 µL of Labeling Master Mix into the tubes containing 50.5 µL of fragmented DNA, giving a total volume of 70 µL.

3. Mix gently the reaction and spìn at 400 x *g* for 1 min at 4 ºC

4. Run the *500K Label* program: 37 ºC, 4 h; 95 ºC, 15 min; hold at 4 ºC.

5. Spin the plate briefly at 400 x *g* to collect the reaction at the bottom of the tube. Samples can be stored at -20 ºC if not proceeding to the next step.

**STEP 9: Target Hybridization**

Before proceeding:

- It is important to alow the arrays to equilibrate to room temperature completely. Unwrap the array and leave on the bench top for 15 minutes.
- DMSO is light sensitive. It should be contained in a dark glass bottle.
- Preparation of the 12X MES Stock: 70.4 g MES Hydrate; 193.3 g MES Sodium salt; 800 mL molecular biology grade water. Mix and adjust volume to 1,000 mL. The pH should be between 6.5 and 6.7. Filter through a 0.2 µm filter.
Do not autoclave. Store between 2 and 8 ºC, and shield from light.

1. Prepare the Hybridization Cocktail Master Mix in the order described. For multiple samples, prepare 5% excess: 12 µL 12X MES; 13 µL DMSO; 13 µL 50X Denhardt's Solution; 3 µL 0.5 M EDTA; 3 µL HSDNA (10 mg/mL); 2 µL OCR 0100; 3 µL Human Cot-1 DNA (1 mg/mL); 1 µL 3% Tween-20; 140 µL 5 M TMACL. Mix well.
2. Transfer each of the labeled samples to a 1.5 mL Eppendorf tube. Aliquot 190 µL of the Hybridization Cocktail Master Mix into the 70 µL of labeled DNA samples, giving a final volume of 260 µL.
3. Heat the 260 µL of hybridization mix and labeled DNA at 99 ºC in a heat block for exactly 10 min to denature.
4. Cool on crushed ice for 10 s.
5. Spin briefly at 400 x *g* in a microfuge to collect any condensate.
6. Place the tubes at 49 ºC for 1 min.
7. Inject 200 µL denatured hybridization cocktail into the array.
8. Hybridize at 49 ºC for 16 to 18 hours at 60 r.p.m. in the oven. The remaining hybridization mix can be stored at -20 ºC for future use.

## Technical notes

The source of RNA is a major determinant of the success for each individual microarray experiment. In this, between 10 to 50 µg of high-quality RNA (usually corresponding to a 50-100 $mm^3$ tumor biopsy) is needed. Ideally, tumor biopsies frozen inmediately after surgical resection in liquid nitrogen (at least at -80 ºC to prevent RNA degradation) should be used.[158,159] This requirement limits the

study of large series of patient samples, most of which are not stored in adequate conditions, especially in retrospective analyses or in series of rare tumors that are collected from different institutions. In addition, this requirement makes problematic to obtain early tumors or biopsies obtained through minimally invasive methods such as fine-needle aspiration.[159-161] An alternative method to preserve biological specimens involves suspending the tissue in a preservative such as RNA*later*, followed by snap freezing of the tissue the next day. This method obviates the inmediate need for nitrogen liquid, preserves the integrity of RNA to be used in microarrays experiments.[160]

Although there are novel methods to extract high-quality RNA from small tumor amounts (even from a single cell) and formalin-fixed tissues, the utility of these RNAs should be extensively evaluated and carefully validated in gene expression microarrays.[160,162]

Tumors are composed of different cell types, including malignant cells, stromal and inflammatory cells and blood vessels. The proportion of these cell populations vary between and within tumors. Because this heterogeneity can complicate the interpretation of microarrays results, a careful selection of the tumors to be included in the study is an important step. In addition, a detailed histopathological analysis of each tumor sample is mandatory. In cases with low percentage of tumor cells, microdissection of the tumor cells in biopsies or cell sorting by flow cytometry in blood, marrow aspirates, effusions or disgregated lymph nodes may be a good choice.[163,164] However, expression of the non-malignant surrounding cells may also be informative, and in some situations the analysis of both isolated tumor cells and whole tumors may be a good choice.[55] One additional issue is the inclusion in the microarray study of normal cell populations to allow the comparisson of tumors genetic profiles with their putative cells of origin and with the normal surrounding cells.[3,165]

In array-based CGH using BAC clones, several factors influence the success of the analysis. First, the general heterogeneity of the spotted BACs, which differ in the proportion of repetitive sequences and gene DNA contents, providing variable signal hybridization intensities. Second, like in gene expression

microarrays, the presence of “contaminating” non-tumoral surrounding cells in the sample. These normal cells have two DNA copies genome-wide (with the exception of X and Y chromosomes in males), and conversely to gene expression profiling, its analysis does not provide any biological information to the study. Thus, array-CGH should limit its application to cases with over 50% of tumoral cells, as lower proportions may yield a normal genomic profile corresponding to the normal cells.[76] Third, the production variability among the different arrays printed on each laboratory, including the few commercially available BAC microarrays. Fourth, these arrays can analyze paraffin-embedded tissues, but it largely depends on the DNA quality and integrity isolated from fixed cells. Despite these difficulties, whole-genome BAC arrays of ~1 Mb resolution (including 3,000 to 4,000 probes) have been succesfully applied to search for genomic changes in many cancer types, allowing an accurate description and mapping of areas of genomic amplification and deletion.[14,120,136,166-168] These alterations can be easily confirmed and visualized in the tumoral cells by complementary fluorescence in situ hybridization (FISH) using the same BACs as probes.[169]

Array-CGH devices can be also applied to scan tumor genomes in other species, predominantly in laboratory mice,[112,170] dogs,[171] and *Drosophila*.[172]

More recently, tiling resolution human DNA microarrays whith over 32,000 overlapping BAC clones covering the entire human genome have been developed, allowing the identification of minute DNA alterations in tumors not previously detected.[109,173] However, the presence of such amount of BAC clones that cannot be individually verified, and the inclusion of only one BAC per array (instead of the 3 to 5 BACs spotted on the 1-Mb BAC arrays) have limited the application of these initial arrays.

One advantage of SNP-CGH arrays is that they only use a test (tumoral) DNA that is hybridized on the chip, without needing any normal DNA as a control. Results of one particular sample are generally “normalized” with respect to available data obtained from the study of a pool of normal DNAs; however, to increase sensitivity and avoid false positive results, the analysis of tumoral and

normal DNAs from each individual in two different arrays is usually recommended.[174-176] Important limitations of this technology include the poor quality results obtained from the analysis of DNAs extracted from paraffin-embedded tissues and the limitation for the detection of areas of UPD in biopsies with more than 50% of non-tumoral cells.

## 8. INTEGRATIVE ONCOGENOMICS: CORRELATION OF GENOMIC ABERRATIONS AND GENE EXPRESSION DATA

We describe step by step our recommended sequence of algorithms and statistical tests to integrate expression data with copy number data.

### 1. Derive gene expression levels and raw copy number data.

The data from the expression and copy number cel files must be pre-processed to remove noise and make the arrays comparable between them.

For gene expression data RMA, GCRMA[177,178], dChip or other methods can be used. The authors recommend the use of RMA, since it has became the "de facto" standard to obtain the expression levels of a gene.

To derive the raw copy number there are several methods such as CNAT, CNAG,[179] dChip, and Aroma.affymetrix.[180] There are marginal differences between them. The most accurate seems to be Aroma.affymetrix. In this case the user has to be confident using R programming language. CNAT, CNAG and dChip provide convenient user interfaces that Aroma does not.

There are other packages for the R programming language (SNPChip). The main disadvantage that occurs in these packages –and not in Aroma.affymetrix– is that all the information of the cel files must be stored in memory limiting the number of arrays to be analyzed to a few tens – depending on the type of array.

There are some special information files related with Affymetrix chips called cdf (*chip definition files*). These files provide the information on how to group each single probe into a set of probes. We recommend using the cdf provided by the Brainarray Website instead of the Affymetrix default files.[181] This website updates the information of these files frequently, improving the results of the analysis. On the other hand, these definition files have the following advantage: a set of probes correspond to a single gene –in the case of Affymetrix a gene can be represented by several set of probes making it difficult to know the correct one.

### 2. Segmentation of the raw copy number data

Copy number alterations occur in segments of the genome –a whole chromosome, an arm of a chromosome or a part of it. This fact can be used to extract the parts –segments- of the genome that have the same copy number.

The procedure to get these parts from the raw copy number data is called segmentation.

There are various algorithms to perform the segmentation. Three of the most widely spread used are CBS (circular binary segmentation),[176] HMM (hidden markov models) and CGHSeq.[182] CNAT, CNAG and dChip provide HMM segmentation whereas Aroma uses CBS segmentation. CHGSeq must be used under the Matlab platform.

A major drawback of HMM is that the number of states –and the corresponding copy numbers– have to be established beforehand. If there is contamination of normal tissue in a tumour sample, copy number will no longer be an integer number, and HMM may fail to discern copy number alterations. CBS and CGHSeq does not have this problem: they provide an estimation of the copy number for each of the segments.

### 3. Assign copy number values to the genes

After the described computations, each SNP have its copy number assigned. These data have to be combined to assign to each gene its corresponding copy number. The copy number for each gene is the mean of the copy number of the segments of the genome where the gene is located.

Genes in which there are copy number changes have to be paid special attention since aberrant splicing forms can occur.

### 4. Remove effects in expression data that are not related with the position in the genome

Segmentation can be also applied to expression data to locate segments of the genome with genes over or under expressed. If applied to expression data, it is better to apply CBS or CGHSeq since there are no obvious means to establish the states beforehand (as needed by HMM). Another possibility is to apply a filter (a moving average across the position in the genome) to the normalized expression data. The weights of this filter can follow a Gaussian distribution (Gaussian filter).

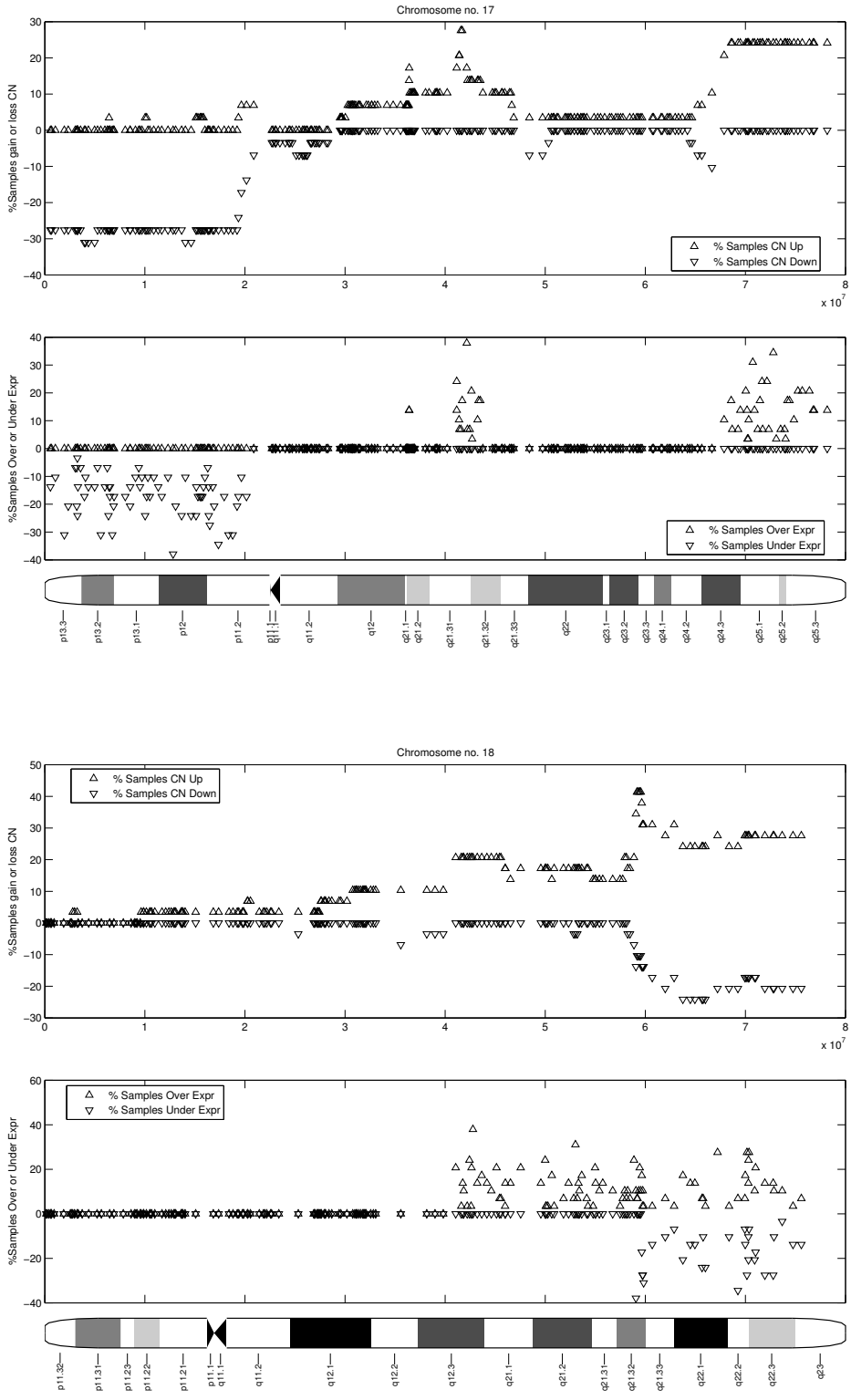


Figure 2. Representation of chromosomes 17 an 18. Top: Percentage of samples with increase or decrease in DNA copy number. Floor: Percentage of samples with overexpressed or underexpressed genes. In the chromosome 18, a high number of samples (more than a 40%) show overexpression of the loci corresponding to the *BCL2* gene.

**5. Detection of cytobands whose genes have their copy number or their expression significantly modified.**

The authors suggest to perform a hypergeometric test to detect which are the cytobands whose genes show a significant variation in copy number (increase or decrease). The hypergeometric test has four parameters: N (the total number of genes), n (the total number of genes in the cytoband), K (the number of genes with copy number up) and k (the number of genes within a cytoband with copy number up). This test provides a pvalue that describes whether the number of genes with copy number up is especially large, i.e. statistically significant, for a particular cytoband. This test can be performed against all the cytobands in the genome (about 300) and for all the samples within the study. The same procedure can be applied to gene expression to detect cytobands cytobands whose genes show a significant variation in their expression.

**6. Global analysis of copy number and expression changes within a study.** A simple procedure to describe which are the locus in the genome that show variation within a study is to show the percentage of samples that have variation in the copy number (up or down) and coherent variation of gene expression, i.e. percentage of samples that shows increase in the copy number and up regulation (Figure 2). This figure shows the results a study with 29 lymphoma cell lines. Two chromosomes are shown (Chr. 17 and 18), each with two graphics. The upper plot shows the percentage of samples with copy number up and down. The lower plot shows the percentage of samples with copy number and expression up (or down). It can be seen, for example, that 17q21.31 shows several genes that have increased both their copy number and their expression. 18q21.31 shows also an increase in copy number and expression. The gene *BCL2* is located in this region.

## REFERENCES


1. Chung CH, Bernard PS, Perou CM. Molecular portraits and the family tree of cancer. Nat Genet. 2002;32 Suppl:533-540.
2. DeRisi J, Penland L, Brown PO, et al. Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat Genet. 1996;14:457-460.
3. Brentani RR, Carraro DM, Verjovski-Almeida S, et al. Gene expression arrays in cancer research: methods and applications. Crit Rev Oncol Hematol. 2005;54:95-105.
4. Staudt LM, Dave S. The biology of human lymphoid malignancies revealed by gene expression profiling. Adv Immunol. 2005;87:163-208.
5. Hoheisel JD. Microarray technology: beyond transcript profiling and genotype analysis. Nat Rev Genet. 2006;7:200-210.
6. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000;403:503-511.
7. Tinker AV, Boussioutas A, Bowtell DD. The challenges of gene expression microarrays for the study of human cancer. Cancer Cell. 2006;9:333-339.
8. Sotiriou C, Piccart MJ. Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? Nat Rev Cancer. 2007;7:545-553.
9. Unger MA, Rishi M, Clemmer VB, et al. Characterization of adjacent breast tumors using oligonucleotide microarrays. Breast Cancer Res. 2001;3:336-341.
10. Kallioniemi A, Kallioniemi OP, Sudar D, et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science. 1992;258:818-821.
11. Pinkel D, Segraves R, Sudar D, et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet. 1998;20:207-211.
12. Solinas-Toldo S, Lampel S, Stilgenbauer S, et al. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. Genes Chromosomes Cancer. 1997;20:399-407.
13. Albertson DG, Ylstra B, Segraves R, et al. Quantitative mapping of amplicon structure by array CGH identifies CYP24 as a candidate oncogene. Nat Genet. 2000;25:144-146.
14. Snijders AM, Nowak N, Segraves R, et al. Assembly of microarrays for genome-wide measurement of DNA copy number. Nat Genet. 2001;29:263-264.
15. Pollack JR, Perou CM, Alizadeh AA, et al. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet. 1999;23:41-46.
16. Barrett MT, Scheffer A, Ben-Dor A, et al. Comparative genomic hybridization using oligonucleotide microarrays and total genomic DNA. Proc Natl Acad Sci U S A. 2004;101:17765-17770.
17. Lindblad-Toh K, Tanenbaum DM, Daly MJ, et al. Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. Nat Biotechnol. 2000;18:1001-1005.
18. Zender L, Lowe SW. Integrative oncogenomic approaches for accelerated cancer-gene discovery. Curr Opin Oncol. 2008;20:72-76.
19. Lowenberg B, Downing JR, Burnett A. Acute myeloid leukemia. N Engl J Med. 1999;341:1051-1062.
20. Rowley JD. Chromosome translocations: dangerous liaisons revisited. Nat Rev Cancer. 2001;1:245-250.

21. Bullinger L, Dohner K, Bair E, et al. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. N Engl J Med. 2004;350:1605-1616.
22. Valk PJ, Verhaak RG, Beijen MA, et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. N Engl J Med. 2004;350:1617-1628.
23. Qian Z, Fernald AA, Godley LA, Larson RA, Le Beau MM. Expression profiling of CD34+ hematopoietic stem/ progenitor cells reveals distinct subtypes of therapy-related acute myeloid leukemia. Proc Natl Acad Sci U S A. 2002;99:14925-14930.
24. Yeoh EJ, Ross ME, Shurtleff SA, et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell. 2002;1:133-143.
25. Ferrando AA, Neuberg DS, Staunton J, et al. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. Cancer Cell. 2002;1:75-87.
26. Kari L, Loboda A, Nebozhyn M, et al. Classification and prediction of survival in patients with the leukemic phase of cutaneous T cell lymphoma. J Exp Med. 2003;197:1477-1488.
27. Thiede C, Steudel C, Mohr B, et al. Analysis of FLT3-activating mutations in 979 patients with acute myelogenous leukemia: association with FAB subtypes and identification of subgroups with poor prognosis. Blood. 2002;99:4326-4335.
28. Armstrong SA, Staunton JE, Silverman LB, et al. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. Nat Genet. 2002;30:41-47.
29. Armstrong SA, Kung AL, Mabon ME, et al. Inhibition of FLT3 in MLL. Validation of a therapeutic target identified by gene expression based classification. Cancer Cell. 2003;3:173-183.
30. Rosenwald A, Wright G, Chan WC, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. N Engl J Med. 2002;346:1937-1947.
31. Hans CP, Weisenburger DD, Greiner TC, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. Blood. 2004;103:275-282.
32. Lossos IS, Czerwinski DK, Alizadeh AA, et al. Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. N Engl J Med. 2004;350:1828-1837.
33. Lossos IS, Morgensztern D. Prognostic biomarkers in diffuse large B-cell lymphoma. J Clin Oncol. 2006;24:995-1007.
34. Lam LT, Davis RE, Pierce J, et al. Small molecule inhibitors of IkappaB kinase are selectively toxic for subgroups of diffuse large B-cell lymphoma defined by gene expression profiling. Clin Cancer Res. 2005;11:28-40.
35. Lam LT, Wright G, Davis RE, et al. Cooperative signaling through the signal transducer and activator of transcription 3 and nuclear factor-{kappa}B pathways in subtypes of diffuse large B-cell lymphoma. Blood. 2008;111:3701-3713.
36. Shipp MA, Ross KN, Tamayo P, et al. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. Nat Med. 2002;8:68-74.
37. Monti S, Savage KJ, Kutok JL, et al. Molecular profiling of diffuse large B-cell lymphoma identifies robust subtypes including one characterized by host inflammatory response. Blood. 2005;105:1851-1861.

38. Chen L, Monti S, Juszczynski P, et al. SYK-dependent tonic B-cell receptor signaling is a rational treatment target in diffuse large B-cell lymphoma. Blood. 2007.
39. Su TT, Guo B, Kawakami Y, et al. PKC-beta controls I kappa B kinase lipid raft recruitment and activation in response to BCR signaling. Nat Immunol. 2002;3:780-786.
40. Smith PG, Wang F, Wilkinson KN, et al. The phosphodiesterase PDE4B limits cAMP-associated PI3K/AKT-dependent apoptosis in diffuse large B-cell lymphoma. Blood. 2005;105:308-316.
41. Robertson MJ, Kahl BS, Vose JM, et al. Phase II study of enzastaurin, a protein kinase C beta inhibitor, in patients with relapsed or refractory diffuse large B-cell lymphoma. J Clin Oncol. 2007;25:1741-1746.
42. Shipp MA. Molecular signatures define new rational treatment targets in large B-cell lymphomas. Hematology Am Soc Hematol Educ Program. 2007;2007:265-269.
43. Polo JM, Dell'Oso T, Ranuncolo SM, et al. Specific peptide interference reveals BCL6 transcriptional and oncogenic mechanisms in B-cell lymphoma cells. Nat Med. 2004;10:1329-1335.
44. Polo JM, Juszczynski P, Monti S, et al. Transcriptional signature with differential expression of BCL6 target genes accurately identifies BCL6-dependent diffuse large B cell lymphomas. Proc Natl Acad Sci U S A. 2007;104:3207-3212.
45. Parekh S, Polo JM, Shaknovich R, et al. BCL6 programs lymphoma cells for survival and differentiation through distinct biochemical mechanisms. Blood. 2007;110:2067-2074.
46. Savage KJ, Monti S, Kutok JL, et al. The molecular signature of mediastinal large B-cell lymphoma differs from that of other diffuse large B-cell lymphomas and shares features with classical Hodgkin lymphoma. Blood. 2003;102:3871-3879.
47. Rosenwald A, Wright G, Leroy K, et al. Molecular diagnosis of primary mediastinal B cell lymphoma identifies a clinically favorable subgroup of diffuse large B cell lymphoma related to Hodgkin lymphoma. J Exp Med. 2003;198:851-862.
48. Kuppers R, Klein U, Schwering I, et al. Identification of Hodgkin and Reed-Sternberg cell-specific genes by gene expression profiling. J Clin Invest. 2003;111:529-537.
49. Klein U, Tu Y, Stolovitzky GA, et al. Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. J Exp Med. 2001;194:1625-1638.
50. Orchard JA, Ibbotson RE, Davis Z, et al. ZAP-70 expression and prognosis in chronic lymphocytic leukaemia. Lancet. 2004;363:105-111.
51. Wiestner A, Rosenwald A, Barry TS, et al. ZAP-70 expression identifies a chronic lymphocytic leukemia subtype with unmutated immunoglobulin genes, inferior clinical outcome, and distinct gene expression profile. Blood. 2003;101:4944-4951.
52. Crespo M, Bosch F, Villamor N, et al. ZAP-70 expression as a surrogate for immunoglobulin-variable-region mutations in chronic lymphocytic leukemia. N Engl J Med. 2003;348:1764-1775.
53. Hummel M, Bentink S, Berger H, et al. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. N Engl J Med. 2006;354:2419-2430.
54. Dave SS, Fu K, Wright GW, et al. Molecular diagnosis of Burkitt's lymphoma. N Engl J Med. 2006;354:2431-2442.
55. Dave SS, Wright G, Tan B, et al. Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells. N Engl J Med. 2004;351:2159-2169.

56. Husson H, Carideo EG, Neuberg D, et al. Gene expression profiling of follicular lymphoma and normal germinal center B cells using cDNA arrays. Blood. 2002;99:282-289.
57. Rosenwald A, Wright G, Wiestner A, et al. The proliferation gene expression signature is a quantitative integrator of oncogenic events that predicts survival in mantle cell lymphoma. Cancer Cell. 2003;3:185-197.
58. Martinez N, Camacho FI, Algara P, et al. The molecular signature of mantle cell lymphoma reveals multiple signals favoring cell survival. Cancer Res. 2003;63:8226-8232.
59. Basso K, Liso A, Tiacci E, et al. Gene expression profiling of hairy cell leukemia reveals a phenotype related to memory B cells with altered expression of chemokine and adhesion receptors. J Exp Med. 2004;199:59-68.
60. Shaughnessy JD, Jr., Zhan F, Burington BE, et al. A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. Blood. 2007;109:2276-2284.
61. Davies FE, Dring AM, Li C, et al. Insights into the multistep transformation of MGUS to myeloma using microarray expression analysis. Blood. 2003;102:4504-4511.
62. Zhan F, Barlogie B, Arzoumanian V, et al. Gene-expression signature of benign monoclonal gammopathy evident in multiple myeloma is linked to good prognosis. Blood. 2007;109:1692-1700.
63. Zhan F, Huang Y, Colla S, et al. The molecular classification of multiple myeloma. Blood. 2006;108:2020-2028.
64. Krivtsov AV, Twomey D, Feng Z, et al. Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. Nature. 2006;442:818-822.
65. Ngo VN, Davis RE, Lamy L, et al. A loss-of-function RNA interference screen for molecular targets in cancer. Nature. 2006;441:106-110.
66. Lenz G, Davis RE, Ngo VN, et al. Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. Science. 2008;319:1676-1679.
67. Peer D, Park EJ, Morishita Y, Carman CV, Shimaoka M. Systemic leukocyte-directed siRNA delivery revealing cyclin D1 as an anti-inflammatory target. Science. 2008;319:627-630.
68. Schlabach MR, Luo J, Solimini NL, et al. Cancer proliferation gene discovery through functional genomics. Science. 2008;319:620-624.
69. Silva JM, Marran K, Parker JS, et al. Profiling essential genes in human mammary cells by multiplex RNAi screening. Science. 2008;319:617-620.
70. Palomero T, Sulis ML, Cortina M, et al. Mutational loss of PTEN induces resistance to NOTCH1 inhibition in T-cell leukemia. Nat Med. 2007;13:1203-1210.
71. Lu J, Getz G, Miska EA, et al. MicroRNA expression profiles classify human cancers. Nature. 2005;435:834-838.
72. Lim LP, Lau NC, Garrett-Engele P, et al. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature. 2005;433:769-773.
73. Calin GA, Dumitru CD, Shimizu M, et al. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A. 2002;99:15524-15529.
74. Calin GA, Liu CG, Sevignani C, et al. MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. Proc Natl Acad Sci U S A. 2004;101:11755-11760.
75. Pekarsky Y, Santanam U, Cimmino A, et al. Tcl1 expression in chronic lymphocytic leukemia is regulated by miR-29 and miR-181. Cancer Res. 2006;66:11590-11593.

76. Pinkel D, Albertson DG. Comparative genomic hybridization. Annu Rev Genomics Hum Genet. 2005;6:331-354.
77. Fukuhara N, Tagawa H, Kameoka Y, et al. Characterization of target genes at the 2p15-16 amplicon in diffuse large B-cell lymphoma. Cancer Sci. 2006;97:499-504.
78. Kasugai Y, Tagawa H, Kameoka Y, Morishima Y, Nakamura S, Seto M. Identification of CCND3 and BYSL as candidate targets for the 6p21 amplification in diffuse large B-cell lymphoma. Clin Cancer Res. 2005;11:8265-8272.
79. Werner CA, Dohner H, Joos S, et al. High-level DNA amplifications are common genetic aberrations in B-cell neoplasms. Am J Pathol. 1997;151:335-342.
80. Bea S, Tort F, Pinyol M, et al. BMI-1 gene amplification and overexpression in hematological malignancies occur mainly in mantle cell lymphomas. Cancer Res. 2001;61:2409-2412.
81. Sanchez-Izquierdo D, Buchonnet G, Siebert R, et al. MALT1 is deregulated by both chromosomal translocation and amplification in B-cell non-Hodgkin lymphoma. Blood. 2003;101:4539-4546.
82. Willis TG, Dyer MJ. The role of immunoglobulin translocations in the pathogenesis of B-cell malignancies. Blood. 2000;96:808-822.
83. Rubio-Moscardo F, Climent J, Siebert R, et al. Mantle-cell lymphoma genotypes identified with CGH to BAC microarrays define a leukemic subgroup of disease and predict patient outcome. Blood. 2005;105:4445-4454.
84. Zhu C, Mills KD, Ferguson DO, et al. Unrepaired DNA breaks in p53-deficient cells lead to oncogenic gene amplification subsequent to translocations. Cell. 2002;109:811-821.
85. Oshiro A, Tagawa H, Ohshima K, et al. Identification of subtype-specific genomic alterations in aggressive adult T-cell leukemia/lymphoma. Blood. 2006;107:4500-4507.
86. He L, Thomson JM, Hemann MT, et al. A microRNA polycistron as a potential human oncogene. Nature. 2005;435:828-833.
87. Tagawa H, Seto M. A microRNA cluster as a target of genomic amplification in malignant lymphoma. Leukemia. 2005;19:2013-2016.
88. Fontana L, Pelosi E, Greco P, et al. MicroRNAs 17-5p-20a-106a control monocytopoiesis through AML1 targeting and M-CSF receptor upregulation. Nat Cell Biol. 2007;9:775-787.
89. Lu Y, Thomson JM, Wong HY, Hammond SM, Hogan BL. Transgenic over-expression of the microRNA miR-17-92 cluster promotes proliferation and inhibits differentiation of lung epithelial progenitor cells. Dev Biol. 2007;310:442-453.
90. Koralov SB, Muljo SA, Galler GR, et al. Dicer ablation affects antibody diversity and cell survival in the B lymphocyte lineage. Cell. 2008;132:860-874.
91. Ventura A, Young AG, Winslow MM, et al. Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters. Cell. 2008;132:875-886.
92. Xiao C, Srinivasan L, Calado DP, et al. Lymphoproliferative disease and autoimmunity in mice with increased miR-17-92 expression in lymphocytes. Nat Immunol. 2008;9:405-414.
93. Dreyling MH, Bullinger L, Ott G, et al. Alterations of the cyclin D1/p16-pRB pathway in mantle cell lymphoma. Cancer Res. 1997;57:4608-4614.
94. Chim CS, Wong KY, Loong F, Lam WW, Srivastava G. Frequent epigenetic inactivation of Rb1 in addition to p15 and p16 in mantle cell and follicular lymphoma. Hum Pathol. 2007;38:1849-1857.

95. Faderl S, Kantarjian HM, Estey E, et al. The prognostic significance of p16(INK4a)/p14(ARF) locus deletion and MDM-2 protein expression in adult acute myelogenous leukemia. Cancer. 2000;89:1976-1982.
96. Gallucci M, Guadagni F, Marzano R, et al. Status of the p53, p16, RB1, and HER-2 genes and chromosomes 3, 7, 9, and 17 in advanced bladder cancer: correlation with adjacent mucosa and pathological parameters. J Clin Pathol. 2005;58:367-371.
97. Kim CH, Yoo JS, Lee CT, et al. FHIT protein enhances paclitaxel-induced apoptosis in lung cancer cells. Int J Cancer. 2006;118:1692-1698.
98. Krug U, Ganser A, Koeffler HP. Tumor suppressor genes in normal and malignant hematopoiesis. Oncogene. 2002;21:3475-3495.
99. Mattioli E, Vogiatzi P, Sun A, et al. Immunohistochemical analysis of pRb2/p130, VEGF, EZH2, p53, p16(INK4A), p27(KIP1), p21(WAF1), Ki-67 expression patterns in gastric cancer. J Cell Physiol. 2007;210:183-191.
100. Mestre-Escorihuela C, Rubio-Moscardo F, Richter JA, et al. Homozygous deletions localize novel tumor suppressor genes in B-cell lymphomas. Blood. 2007;109:271-280.
101. Tagawa H, Karnan S, Suzuki R, et al. Genome-wide array-based CGH for mantle cell lymphoma: identification of homozygous deletions of the proapoptotic gene BIM. Oncogene. 2005;24:1348-1358.
102. Pasqualucci L, Compagno M, Houldsworth J, et al. Inactivation of the PRDM1/BLIMP1 gene in diffuse large B cell lymphoma. J Exp Med. 2006;203:311-317.
103. Tam W, Gomez M, Chadburn A, Lee JW, Chan WC, Knowles DM. Mutational analysis of PRDM1 indicates a tumor-suppressor role in diffuse large B-cell lymphomas. Blood. 2006;107:4090-4100.
104. Ross CW, Ouillette PD, Saddler CM, Shedden KA, Malek SN. Comprehensive analysis of copy number and allele status identifies multiple chromosome defects underlying follicular lymphoma pathogenesis. Clin Cancer Res. 2007;13:4777-4785.
105. Fitzgibbon J, Iqbal S, Davies A, et al. Genome-wide detection of recurring sites of uniparental disomy in follicular and transformed follicular lymphoma. Leukemia. 2007;21:1514-1520.
106. Honma K, Tsuzuki S, Nakagawa M, et al. TNFAIP3 is the target gene of chromosome band 6q23.3-q24.1 loss in ocular adnexal marginal zone B cell lymphoma. Genes Chromosomes Cancer. 2008;47:1-7.
107. Kim WS, Honma K, Karnan S, et al. Genome-wide array-based comparative genomic hybridization of ocular marginal zone B cell lymphoma: comparison with pulmonary and nodal marginal zone B cell lymphoma. Genes Chromosomes Cancer. 2007;46:776-783.
108. Thelander EF, Ichimura K, Corcoran M, et al. Characterization of 6q deletions in mature B cell lymphomas and childhood acute lymphoblastic leukemia. Leuk Lymphoma. 2008;49:477-487.
109. de Leeuw RJ, Davies JJ, Rosenwald A, et al. Comprehensive whole genome array CGH profiling of mantle cell lymphoma model genomes. Hum Mol Genet. 2004;13:1827-1837.
110. Hodgson G, Hager JH, Volik S, et al. Genome scanning with array CGH delineates regional alterations in mouse islet carcinomas. Nat Genet. 2001;29:459-464.
111. Mao JH, Perez-Losada J, Wu D, et al. Fbxw7/Cdc4 is a p53-dependent, haploinsufficient tumour suppressor gene. Nature. 2004;432:775-779.

112. Snijders AM, Nowak NJ, Huey B, et al. Mapping segmental and sequence variations among laboratory mice using BAC array CGH. Genome Res. 2005;15:302-311.
113. Mullighan CG, Goorha S, Radtke I, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature. 2007;446:758-764.
114. Cobaleda C, Jochum W, Busslinger M. Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. Nature. 2007;449:473-477.
115. Xie H, Ye M, Feng R, Graf T. Stepwise reprogramming of B cells into macrophages. Cell. 2004;117:663-676.
116. Akasaka T, Balasas T, Russell LJ, et al. Five members of the CEBP transcription factor family are targeted by recurrent IGH translocations in B-cell precursor acute lymphoblastic leukemia (BCP-ALL). Blood. 2007;109:3451-3461.
117. Dohner H, Stilgenbauer S, Benner A, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med. 2000;343:1910-1916.
118. Zenz T, Dohner H, Stilgenbauer S. Genetics and risk-stratified approach to therapy in chronic lymphocytic leukemia. Best Pract Res Clin Haematol. 2007;20:439-453.
119. Schwaenen C, Nessling M, Wessendorf S, et al. Automated array-based genomic profiling in chronic lymphocytic leukemia: development of a clinical tool and discovery of recurrent genomic alterations. Proc Natl Acad Sci U S A. 2004;101:1039-1044.
120. Kohlhammer H, Schwaenen C, Wessendorf S, et al. Genomic DNA-chip hybridization in t(11;14)-positive mantle cell lymphomas shows a high frequency of aberrations and allows a refined characterization of consensus regions. Blood. 2004;104:795-801.
121. Bea S, Ribas M, Hernandez JM, et al. Increased number of chromosomal imbalances and high-level DNA amplifications in mantle cell lymphoma are associated with blastoid variants. Blood. 1999;93:4365-4374.
122. Salaverria I, Zettl A, Bea S, et al. Specific secondary genetic alterations in mantle cell lymphoma provide prognostic information independent of the gene expression-based proliferation signature. J Clin Oncol. 2007;25:1216-1222.
123. Saddler C, Ouillette P, Kujawski L, et al. Comprehensive biomarker and genomic analysis identifies P53 status as the major determinant of response to MDM2 inhibitors in chronic lymphcytic leukemia. Blood. 2007.
124. Raghavan M, Lillington DM, Skoulakis S, et al. Genome-wide single nucleotide polymorphism analysis reveals frequent partial uniparental disomy due to somatic recombination in acute myeloid leukemias. Cancer Res. 2005;65:375-378.
125. Fitzgibbon J, Smith LL, Raghavan M, et al. Association between acquired uniparental disomy and homozygous gene mutation in acute myeloid leukemias. Cancer Res. 2005;65:9152-9154.
126. Baxter EJ, Scott LM, Campbell PJ, et al. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. Lancet. 2005;365:1054-1061.
127. Kralovics R, Passamonti F, Buser AS, et al. A gain-of-function mutation of JAK2 in myeloproliferative disorders. N Engl J Med. 2005;352:1779-1790.
128. Levine RL, Wadleigh M, Cools J, et al. Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. Cancer Cell. 2005;7:387-397.
129. Jones AV, Kreil S, Zoi K, et al. Widespread occurrence of the JAK2 V617F mutation in chronic myeloproliferative disorders. Blood. 2005;106:2162-2168.
130. Flotho C, Steinemann D, Mullighan CG, et al. Genome-wide single-nucleotide polymorphism analysis in juvenile myelomonocytic leukemia identifies uniparental

disomy surrounding the NF1 locus in cases associated with neurofibromatosis but not in cases with mutant RAS or PTPN11. Oncogene. 2007;26:5816-5821.
131. Nielaender I, Martin-Subero JI, Wagner F, Martinez-Climent JA, Siebert R. Partial uniparental disomy: a recurrent genetic mechanism alternative to chromosomal deletion in malignant lymphoma. Leukemia. 2006;20:904-905.
132. Sellick GS, Goldin LR, Wild RW, et al. A high-density SNP genome-wide linkage search of 206 families identifies susceptibility loci for chronic lymphocytic leukemia. Blood. 2007;110:3326-3333.
133. Lockhart DJ, Winzeler EA. Genomics, gene expression and DNA arrays. Nature. 2000;405:827-836.
134. Pollack JR, Sorlie T, Perou CM, et al. Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. Proc Natl Acad Sci U S A. 2002;99:12963-12968.
135. Hyman E, Kauraniemi P, Hautaniemi S, et al. Impact of DNA amplification on gene expression patterns in breast cancer. Cancer Res. 2002;62:6240-6245.
136. Martinez-Climent JA, Alizadeh AA, Segraves R, et al. Transformation of follicular lymphoma to diffuse large cell lymphoma is associated with a heterogeneous set of DNA copy number and gene expression alterations. Blood. 2003;101:3109-3117.
137. Myers CL, Dunham MJ, Kung SY, Troyanskaya OG. Accurate detection of aneuploidies in array CGH and gene expression microarray data. Bioinformatics. 2004;20:3533-3543.
138. Yi Y, Mirosevich J, Shyr Y, Matusik R, George AL, Jr. Coupled analysis of gene expression and chromosomal location. Genomics. 2005;85:401-412.
139. La Rosa P, Viara E, Hupe P, et al. VAMP: visualization and analysis of array-CGH, transcriptome and other molecular profiles. Bioinformatics. 2006;22:2066-2073.
140. Margolin AA, Nemenman I, Basso K, et al. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. BMC Bioinformatics. 2006;7 Suppl 1:S7.
141. Garraway LA, Widlund HR, Rubin MA, et al. Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. Nature. 2005;436:117-122.
142. Yu J, Cao Q, Mehra R, et al. Integrative genomics analysis reveals silencing of beta-adrenergic signaling by polycomb in prostate cancer. Cancer Cell. 2007;12:419-431.
143. Overholtzer M, Zhang J, Smolen GA, et al. Transforming properties of YAP, a candidate oncogene on the chromosome 11q22 amplicon. Proc Natl Acad Sci U S A. 2006;103:12405-12410.
144. Zender L, Spector MS, Xue W, et al. Identification and validation of oncogenes in liver cancer using an integrative oncogenomic approach. Cell. 2006;125:1253-1267.
145. Kim M, Gans JD, Nogueira C, et al. Comparative oncogenomics identifies NEDD9 as a melanoma metastasis gene. Cell. 2006;125:1269-1281.
146. Chang TC, Yu D, Lee YS, et al. Widespread microRNA repression by Myc contributes to tumorigenesis. Nat Genet. 2008;40:43-50.
147. Martinez-Climent JA, Vizcarra E, Sanchez D, et al. Loss of a novel tumor suppressor gene locus at chromosome 8p is associated with leukemic mantle cell lymphoma. Blood. 2001;98:3479-3482.
148. Rubio-Moscardo F, Blesa D, Mestre C, et al. Characterization of 8p21.3 chromosomal deletions in B-cell lymphoma: TRAIL-R1 and TRAIL-R2 as candidate dosage-dependent tumor suppressor genes. Blood. 2005;106:3214-3222.

149. Ramalingam A, Duhadaway JB, Sutanto-Ward E, et al. Bin3 deletion causes cataracts and increased susceptibility to lymphoma during aging. Cancer Res. 2008;68:1683-1690.
150. Takeyama K, Monti S, Manis JP, et al. Integrative analysis reveals 53BP1 copy loss and decreased expression in a subset of human diffuse large B-cell lymphomas. Oncogene. 2008;27:318-322.
151. Huusko P, Ponciano-Jackson D, Wolf M, et al. Nonsense-mediated decay microarray analysis identifies mutations of EPHB2 in human prostate cancer. Nat Genet. 2004;36:979-983.
152. Zardo G, Tiirikainen MI, Hong C, et al. Integrated genomic and epigenomic analyses pinpoint biallelic gene inactivation in tumors. Nat Genet. 2002;32:453-458.
153. Stransky N, Vallot C, Reyal F, et al. Regional copy number-independent deregulation of transcription in cancer. Nat Genet. 2006;38:1386-1396.
154. Saito Y, Liang G, Egger G, et al. Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. Cancer Cell. 2006;9:435-443.
155. Carrasco DR, Tonon G, Huang Y, et al. High-resolution genomic profiles define distinct clinico-pathogenetic subgroups of multiple myeloma patients. Cancer Cell. 2006;9:313-325.
156. Keats JJ, Fonseca R, Chesi M, et al. Promiscuous mutations activate the noncanonical NF-kappaB pathway in multiple myeloma. Cancer Cell. 2007;12:131-144.
157. Annunziata CM, Davis RE, Demchenko Y, et al. Frequent engagement of the classical and alternative NF-kappaB pathways by diverse genetic abnormalities in multiple myeloma. Cancer Cell. 2007;12:115-130.
158. Chowdary D, Lathrop J, Skelton J, et al. Prognostic gene expression signatures can be measured in tissues collected in RNAlater preservative. J Mol Diagn. 2006;8:31-39.
159. Wang E, Miller LD, Ohnmacht GA, Liu ET, Marincola FM. High-fidelity mRNA amplification for gene profiling. Nat Biotechnol. 2000;18:457-459.
160. Mazumder A, Wang Y. Gene-expression signatures in oncology diagnostics. Pharmacogenomics. 2006;7:1167-1173.
161. Florell SR, Coffin CM, Holden JA, et al. Preservation of RNA for functional genomic studies: a multidisciplinary tumor bank protocol. Mod Pathol. 2001;14:116-128.
162. Chen J, Byrne GE, Jr., Lossos IS. Optimization of RNA extraction from formalin-fixed, paraffin-embedded lymphoid tissues. Diagn Mol Pathol. 2007;16:61-72.
163. Barrett MT, Glogovac J, Prevo LJ, Reid BJ, Porter P, Rabinovitch PS. High-quality RNA and DNA from flow cytometrically sorted human epithelial cells and tissues. Biotechniques. 2002;32:888-890, 892, 894, 896.
164. Aoyagi K, Tatsuta T, Nishigaki M, et al. A faithful method for PCR-mediated global mRNA amplification and its integration into microarray analysis on laser-captured cells. Biochem Biophys Res Commun. 2003;300:915-920.
165. Alizadeh A, Eisen M, Davis RE, et al. The lymphochip: a specialized cDNA microarray for the genomic-scale analysis of gene expression in normal and malignant lymphocytes. Cold Spring Harb Symp Quant Biol. 1999;64:71-78.
166. Greshock J, Naylor TL, Margolin A, et al. 1-Mb resolution array-based comparative genomic hybridization using a BAC clone set optimized for cancer gene analysis. Genome Res. 2004;14:179-187.

167. Tagawa H, Suguro M, Tsuzuki S, et al. Comparison of genome profiles for identification of distinct subgroups of diffuse large B-cell lymphoma. Blood. 2005;106:1770-1777.
168. Fiegler H, Carr P, Douglas EJ, et al. DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. Genes Chromosomes Cancer. 2003;36:361-374.
169. Kallioniemi A, Visakorpi T, Karhu R, Pinkel D, Kallioniemi OP. Gene Copy Number Analysis by Fluorescence in Situ Hybridization and Comparative Genomic Hybridization. Methods. 1996;9:113-121.
170. Chung YJ, Jonkers J, Kitson H, et al. A whole-genome mouse BAC microarray with 1-Mb resolution for analysis of DNA copy number changes by array comparative genomic hybridization. Genome Res. 2004;14:188-196.
171. Thomas R, Fiegler H, Ostrander EA, Galibert F, Carter NP, Breen M. A canine cancer-gene microarray for CGH analysis of tumors. Cytogenet Genome Res. 2003;102:254-260.
172. Fan C, Long M. A new retroposed gene in Drosophila heterochromatin detected by microarray-based comparative genomic hybridization. J Mol Evol. 2007;64:272-283.
173. Ishkanian AS, Malloff CA, Watson SK, et al. A tiling resolution DNA microarray with complete coverage of the human genome. Nat Genet. 2004;36:299-303.
174. Lin M, Wei LJ, Sellers WR, Lieberfarb M, Wong WH, Li C. dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. Bioinformatics. 2004;20:1233-1240.
175. McCarroll SA, Hadnott TN, Perry GH, et al. Common deletion polymorphisms in the human genome. Nat Genet. 2006;38:86-92.
176. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics. 2004;5:557-572.
177. Cope LM, Irizarry RA, Jaffee HA, Wu Z, Speed TP. A benchmark for Affymetrix GeneChip expression measures. Bioinformatics. 2004;20:323-331.
178. Wu Z, Irizarry RA. Preprocessing of oligonucleotide array data. Nat Biotechnol. 2004;22:656-658; author reply 658.
179. Nannya Y, Sanada M, Nakazaki K, et al. A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. Cancer Res. 2005;65:6071-6079.
180. Bengtsson H, Irizarry R, Carvalho B, Speed TP. Estimation and assessment of raw copy numbers at the single locus level. Bioinformatics. 2008;24:759-767.
181. Dai M, Wang P, Boyd AD, et al. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. Nucleic Acids Res. 2005;33:e175.
182. Lai WR, Johnson MD, Kucherlapati R, Park PJ. Comparative analysis of algorithms for identifying amplifications and deletions in array CGH data. Bioinformatics. 2005;21:3763-3770.