

Est.  
1841

YORK  
ST JOHN  
UNIVERSITY

Baumgartner, Adi ORCID logoORCID:  
<https://orcid.org/0000-0001-7042-0308>, Hartleb, Veronika and  
Taylor, Jim ORCID logoORCID: <https://orcid.org/0000-0003-2758-941X> (2019) Comparative Genomic Hybridization (CGH) in  
Genotoxicology. In: Dhawan, Alok and Bajpayee, Mahima, (eds.)  
Protocols in Genotoxicity Assessment. 2nd ed. Springer, pp. 209-  
268

Downloaded from: <https://ray.yorks.ac.uk/id/eprint/3659/>

The version presented here may differ from the published version or version of record. If  
you intend to cite from the work you are advised to consult the publisher's version:

[https://link.springer.com/protocol/10.1007/978-1-4939-9646-9\\_11](https://link.springer.com/protocol/10.1007/978-1-4939-9646-9_11)

Research at York St John (RaY) is an institutional repository. It supports the principles of  
open access by making the research outputs of the University available in digital form.  
Copyright of the items stored in RaY reside with the authors and/or other copyright  
owners. Users may access full text items free of charge, and may download a copy for  
private study or non-commercial research. For further reuse terms, see licence terms  
governing individual outputs. [Institutional Repository Policy Statement](#)

# RaY

Research at the University of York St John

For more information please contact RaY at [ray@yorks.ac.uk](mailto:ray@yorks.ac.uk)

Chapter XX

## **Comparative Genomic Hybridization (CGH) in Genotoxicology**

**Adi Baumgartner<sup>1</sup>, Veronika Hartleb<sup>2</sup> and Jim D. Taylor<sup>1</sup>**

<sup>1</sup>York St John University, School of Health Sciences, Biomedical Science, Lord Mayor's Walk, York, YO31 7EX, United Kingdom

<sup>2</sup>MRC Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford, OX3 9DS, United Kingdom

Email: [a.baumgartner@yorks.ac.uk](mailto:a.baumgartner@yorks.ac.uk)

Phone: +44-1904-87-6714

## **Abstract**

In the past two decades, comparative genomic hybridization (CGH) and array CGH have become indispensable tools in clinical diagnostics and toxicological risk assessment. Initially developed for the genome-wide screening of chromosomal imbalances, i.e., copy-number variations in tumor cells, both CGH and array CGH have been employed in genotoxicology and most recently in toxicogenomics. The latter allows a multi-endpoint analysis of how particular genes react to toxic agents, revealing changes in signaling pathways and other underlying molecular mechanisms. This chapter provides background on the use of CGH and array CGH in the context of genotoxicology, and also a protocol for conventional CGH, so that the basic principles of this methodology can be better understood. Conventional and array CGH investigate DNA expression patterns, copy-number variations across the whole genome, and loss of heterozygosity after genotoxic damage. Array CGH is still cost intensive, but produces exponentially more data, requiring suitable analytical algorithms and sophisticated bioinformatic analysis. As toxicogenomics is an emerging sub-discipline of toxicology research, effectively evaluating toxicogenomic microarray data can be hugely advantageous for human risk assessment, even though international regulatory guidelines on toxicogenomics have yet to be fully agreed and implemented.

## **Key Words**

Comparative genomic hybridization, CGH, array CGH, microarray, genomic imbalances, genotoxicology, toxicogenomics

## **Running head**

CGH in Genotoxicology

## 1. Introduction

For cytogenetics and clinical diagnostics, the introduction of a highly versatile molecular cytogenetic technique named fluorescence *in situ* hybridization (FISH) in the mid-1980s was revolutionary for the evaluation of cytogenetic aberrations and abnormalities at the molecular level (*1*). Since then, this technique has been refined and adapted countless times (*2*). The plethora of applications being used range from fiber FISH (*3*), multiplex FISH (*4*) and spectral karyotyping (*5*), combined binary ratio labeling (COBRA) FISH (*6*) to 3D FISH that have been established to study interphase nuclear chromosome compartments (*7,8*).

Continuing the success of FISH, more recently Cas9-mediated FISH (CASFISH) has been developed for living and fixed cells, but unlike the aforementioned, this methodology does not use nucleic acid probes (*9,10*). Depending on the assay in use, the resolution varies from 5kb to 5Mb when examining fluorescent signals at the level of chromatin strands, at interphases or metaphases (*11*). With the exception of CASFISH, the above assays require complementary probes or probe sets, differentially labeled with one or more fluorescent dyes, to be hybridized to the target DNA being investigated, e.g., gene loci on metaphase chromosomes or in interphase nuclei. Changes within the target DNA can then be visualized and evaluated. However, it has proved extremely difficult and challenging to prepare metaphase spreads from certain cells or tissues such as solid tumors. Thus, comparative genomic hybridization (CGH) was developed by modification of the extensively used multicolor FISH assays, to address the limitations with solid tumors. In the early 1990s, virtually at the same time, two groups, one in the USA at the University of California in San Francisco (*12*) and the other group in Heidelberg, Germany (*13*), recognized the principles of CGH which represented a significant improvement over other fluorescence-based *in situ*

hybridization techniques, mainly because there was no need to culture cells from the tissue of interest.

### ***1.1. Principle of Comparative Genomic Hybridization (CGH)***

Comparative genomic hybridization permits the detection of unbalanced copy number variations (CNV) (**14**) by using differentially labeled genomic DNA from a tissue of interest (test sample) and a reference DNA (control sample). Depending on the application, such investigations can be undertaken at the whole genome level; or targeted to specific chromosomes or chromosomal regions. The reference DNA is usually labeled with a green fluorochrome and mixed in equal parts with the test DNA labeled with a red fluorochrome. Both DNA samples are then simultaneously hybridized to a matrix, e.g., metaphase chromosomes or oligos, where the labeled fragments compete for hybridization to their specific loci. Fluorescence microscopy and CGH-specific computer-based analysis algorithms are then employed to detect the fluorescence profiles for each chromosome or chromosomal area of interest. For unaffected genomic areas, the ratio of hybridized test and control DNA is approximately equal and is thus seen as the additive color, yellow. Loss of genetic material in the test sample will result in quantitatively less green-labeled DNA hybridization and thus the color of this locus will shift towards red. Similarly, additional genetic material in the test sample will cause a green-shift in the observed fluorescence profile. The resulting intensity ratios for both colors can then be computed. Hence, in a perfect (but theoretical) situation without data normalization or measurement errors, where all tumor cells have identical genomic alterations and are uncontaminated by cells from surrounding normal tissue, the normal copy-neutral ratio would correspond to  $\log_2 \frac{2}{2} = 0$  because the reference and test DNA fragments both have two copies and are present in equal

measure. Thus, no change in copy number would be inferred from the resulting profile. The  $\log_2$ -transformed mean intensity ratios of single-copy loss (e.g., of a chromosomal region) would be calculated as  $\log_2 \frac{1}{2} = -1$ , while a single-copy gain would result in  $\log_2 \frac{3}{2} = 0.58$ . Multiple-copy gains, i.e., gene amplifications, frequently found for oncogenes in tumors would relate to  $\log_2 \frac{4}{2} = 1$ ,  $\log_2 \frac{5}{2} = 1.3$  and so on. On the other hand, loss or deletion of both copies of a gene, often associated with tumor-suppressor mutations, would correspond to a negative infinite value ( $-\infty$ ). In this hypothetical situation, the genomic alterations can be easily deduced from the data without the use of complex statistical techniques.

## ***1.2. Chromosomal CGH***

The conventional chromosomal CGH method was developed exclusively to compare whole genomes for copy number variations (CNV), e.g., gain and/or loss of chromosomal DNA sequences. The major difference between CGH and the commonly employed FISH technique is that metaphase chromosomes serve only as a hybridization matrix for comparative genomic hybridization and do not themselves represent target structures to be analyzed. Metaphase chromosome spreads on glass slides originate from cultures of stimulated lymphocytes from healthy individuals and can either be self-prepared or obtained commercially. The quality of the chromosome spreads, the degree of chromosome condensation and the density of the spreads on the slide are important criteria for the selection of these slides for CGH since they affect the detection resolution. Genomic DNA is usually isolated from a healthy individual (referred to as reference or control DNA) and from specific target cells/tissue (commonly referred to as test DNA), e.g., tissue from a solid tumor (*15*). These DNAs are each labeled with different fluorochromes, for instance the control DNA in green, e.g., using fluorescein isothiocyanate (FITC), and the test DNA in red, e.g., using Cy3 or rhodamine, and then co-

hybridized onto the chromosomal matrix (metaphase chromosomes). After hybridization, the ratio of fluorescence intensities at each locus displays the cytogenetic information regarding genomic changes in the test genome, i.e. CNV, with respect to the control DNA used (**Figure 1**).

The main advantage of chromosomal CGH is that there is no need to culture and prepare metaphase spreads from target cells such as tumor cells in order to analyze CNV or numerical abnormalities in the genome of the studied cells. The lowest detection limit of strongly over-expressed DNA in conventional CGH was found to be 0.25 Mb. The maximum resolution after losing one homologue was approximately 2 Mb, while the loss of both homologue chromosomes resulted in a resolution of 1 Mb. The resolution for amplified or deleted regions might practically range from 5-10 Mb (*16,17*).

**Section 3** of this chapter shows a standard protocol for chromosomal CGH using directly labeled fluorescent DNA probes, prepared using nick translation.

### **1.2.1. Using Chromosomal CGH in Genotoxicology**

Initially developed to investigate chromosomal changes in the genome of solid tumors, chromosomal CGH has also been employed to study the impact of genotoxins on the genome. As a screening tool for mutagenicity testing of chemicals, Corso and Parry (*18*) developed the cell line MCL-5 for use with the CGH assay by transfecting the human B cell-derived lymphoblastoid cell line AHH-1 TK<sup>+/-</sup> to become metabolically active and to express the cytochrome P450 enzymes CYP1A2, CYP2A6, CYP2E1, and CYP3A4 but also microsomal epoxide hydrolase (*18*). Subsequently, toxicological studies have been undertaken to determine CNV within the whole genome using various cell lines such as K562, MCF-7 and MCF-10A or by using animal tumor cells (e.g., rat gastric tumor cells). In particular,



resistance to various cytostatic drugs (19), to the gastric tumor inducer and alkylating agent N-methyl-N'-nitrosoguanidine (MNNG) (20), to xenoestrogens (21) and to the soy isoflavone genistein (22) have been evaluated. An interesting approach to evaluate environmental toxins such as mycotoxins and viruses in relation to their carcinogenicity was carried out by Wong and colleagues (23). Using cells from human hepatocarcinomas from different geographic locations around the globe, a correlation between different risk factors such as aflatoxin intoxication or hepatitis virus infection and hepatocarcinoma prevalence was shown (23). Subsequently, patterns of chromosomal gains and losses were successfully determined by chromosomal CGH. One finding indicated that hepatitis C virus-related samples from Japan had a characteristically high incidence of an 11q13 gain in the tumor's genome (23). With the completion of the Human Genome Project, sequence information became publicly available, revolutionizing biomedical research and allowing genome wide investigation using microarray technology (24). Consequently, the nowadays more common array CGH, replaced the classical chromosomal CGH in research; however, chromosomal CGH still has potential for cost-effectively and rapidly assessing CNV.

### ***1.3. Array CGH***

In the late 1990s, the chromosomal CGH method was refined (25) using microarrays of genomic DNA fragments instead of whole metaphase chromosomes as hybridization targets (**Figure 2**). This development significantly increased the resolution of the methodology. Early assays employed arrays with large human DNA inserts in plasmids like bacterial artificial chromosomes (BAC), P1-derived artificial chromosomes (PAC) or yeast artificial chromosomes (YAC) and other sources of well characterized DNA sequences, e.g., cosmids or sub-cloned cDNA. The physical length of a single BAC spotted on such an array can be up

to 200 kb long. These so-called BAC-arrays reached 3,000 to 30,000 spots, but the resolution was found to be rather limited at around 3 Mb (26).

Alongside the development of BAC-arrays, oligonucleotide-based microarrays were introduced with elements of 25 to 85-mer high-density synthetic oligonucleotides or expressed sequence tags (EST) on up to 250,000 printed spots providing a higher resolution of 50-100 kb (27,28). In 2005, Davies and his team illustrated the evolution and diversity of array CGH and its application to cancer genomes (29).

On such an array, normal (control) fluorochrome-conjugated oligonucleotides are covalently bound to the surface of a glass slide, providing a much higher resolution hybridization matrix than the metaphase spread used in chromosomal CGH. After hybridization of the test sample (prepared and labeled as per chromosomal CGH) with the array, the ratio of intensities of both fluorescent dyes can be visualized and subsequently evaluated. The amount of fluorescence conjugate bound to each microarray spot corresponds to the level of the particular gene(s) expressed in the test specimen (30,28). At the dawn of oligonucleotide arrays, a feasibility study showed that this technology could only provide a facile overview of gene expression responses relevant to drug metabolism and toxicology (31). But, with the increase in numbers of mapped oligonucleotides, the resolution and efficiency has increased, leading to further refinements and variants of this array-based method. The exon array CGH assay, for instance, evaluates exogenic CNV, hence, only targeting exons within genes (32,33). Although primarily developed for pre-implantation genetic diagnosis to assess numerical abnormalities on single blastomeres, single-cell array CGH assay can also be employed to focus on the genotoxic consequences of toxicant exposure at the single cell level (34). In this variant of array CGH, only one specific cell, e.g., a blastomere, is analyzed for its genomic CNV (35,36).

The disadvantage of the above array CGH method is that since it is based on the use of chromosomal DNA fragments as target structures, no information on zygosity can be obtained – this also applies to conventional chromosomal CGH. Thus, manufacturers of array-based chips began to add single nucleotide polymorphism (SNP) containing nucleic acids sequences (25 bp long with a centered SNP) together with oligonucleotides onto microarrays. Moreover, straight SNP arrays were able to assess DNA sequence variations within chromosomes, individuals or even species (37). Besides distinguishing loss of heterozygosity events, these high-density SNP arrays also proved to be highly efficient for evaluating genome-wide CNV such as genomic amplifications or homozygous deletions (38,39). State-of-the-art high resolution CGH arrays (DNA chips) are nowadays capable of detecting and assessing segmental DNA CNV at a kilobase pair resolution (spatial resolution of around 35 kb) with more than 900,000 SNPs and close to one million probes, e.g., the Genome-Wide Human SNP Array 6.0 from Affymetrix (40,41). Nevertheless, despite the growing number of very sophisticated microarray DNA chips, common BAC arrays with a spatial resolution in the 150 kb range, exhibit the best signal-to-noise ratio when compared with oligonucleotide array platforms and might be better suited for the detection of single-copy aberrations (42).

### **1.3.1. Using array CGH in Genotoxicology**

For genotoxicity evaluation and toxicogenomics, arrays using bacterial artificial chromosomes or oligonucleotides, have been employed using mainly cell lines as target cells for toxicants. By using BAC arrays containing 6,500 BAC clones representing 0.5 Mb genomic resolution, Herzog and colleagues evaluated CNV and chromosomal instability in

mouse lung adenocarcinoma cells induced by the potent human lung carcinogen nicotine-derived nitrosamine ketone NNK (4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone) present in tobacco smoke. Significantly, more gross chromosomal changes were found in NNK-induced tumors compared with spontaneous tumors (43).

Oligonucleotide arrays have been employed to evaluate genotoxins as well. Usually, cell lines but also human and animal tumor cells derived from primary cells are the cells of choice to study the impact of genotoxins on the DNA, genes and underlying regulatory/signaling pathways. At the end of the 1990s, the National Institute of Environmental Health Sciences (NIEHS) Microarray Center developed a so-called ToxChip with spotted human cDNAs containing clusters of up to 12,000 different cloned genes (44). This array CGH chip technology allowed the screening and classification of toxicants by their gene expression pattern and therefore helped to elucidate their mechanism of action. This permitted the evaluation of unknown compounds by comparison with the signatures of well-known toxicants. Heinloth *et al.* (45,46) used these chips to assess the effect of 5 Gy  $\gamma$ -radiation, 7.5 J/m<sup>2</sup> UV-radiation and oxidative stress (75  $\mu$ M *tert*-butyl hydroperoxide) on dermal fibroblasts. The three treatments resulted in distinct patterns, indicating an involvement of ATM in regulation of transcription factors such as SP1, AP1 and MTF1 (46) as well as cyclin E-associated kinase activity reduction (45). Similarly, Brown *et al.* (47) used custom-printed mouse chromosome-2 arrays with unprecedented 148-bp resolution to examine the effects of radiation quality on chromosome-2 deletions leading to radiation-induced acute myeloid leukemia (rAML). The study revealed a complex relationship between rAML and chromosome-2 aberrations, but no difference between chromosome-2 deletions following low LET or high LET X-rays/neutron exposure (47).

In order to detect small copy-number changes, human breast cancer MCF-10A cells were treated with the isoflavone genistein for three months and then evaluated using conventional

chromosomal CGH together with high-density oligonucleotide microarray CGH. Besides finding a characteristic deletion on 9p21, long-term exposure to genistein might increase chromosomal imbalance (22). Microarray analysis has proved to be a valuable tool to assess CNV and the potential cancer risk also with regards to cigarette smoking (48) and tobacco carcinogens (49). It has also allowed the evaluation of gene expression changes of hepatocarcinogens such as the mycotoxins aflatoxin (50) and citrinin (51). In addition, Huang *et al.* (52) used a microarray expression and genotyping assay to assess gene expression and genotypes (heterozygous or homozygous) in the human mammary gland MCF-10F cell-line after treatment with the steroid hormone 17 $\beta$ -estradiol (E<sub>2</sub>), which may induce complete neoplastic transformation of this epithelial cell line (52).

It is clear that despite the development of other molecular assays to detect CNV, array CGH technology continues to evolve. Reassuringly, a comparison of CGH technologies across microarray platforms such as BAC array, genotyping oligonucleotide array and RNA expression microarray carried out by Hester *et al.* showed high concordance among CNV detection in all platforms, suggesting that the chosen CGH technology may depend more on practical considerations than differences in assay performance (42).

### **1.3.2. Array CGH Evaluation**

Microarray technology is a particularly powerful tool for identifying new genotoxic substances and their mechanisms of action. It also allows the determination of different susceptibility thresholds for a range of tissues and cell types (53). However, careful planning and choice of analysis methodology is required for optimal results (54). Differences between array types (e.g., genomic or expression arrays) and parameters such as; signal-to-noise

ratios; low standard deviations (SD) of intensity ratios as well as an optimized cot-1 DNA to probe DNA ratio must be considered for successful implementation of this technology (55). A high-throughput method like array CGH can generate a huge amount of data regarding CNV across the whole genome, especially given the ever-increasing resolution of the array chips. Commonly, array CGH fluorescence intensity ratios for each probe are transformed onto a binary logarithmic ( $\log_2$ ) scale to show proportional change in copy number and provide the most suitable information about genome-wide changes. However, under real world conditions, the  $\log_2$  values may vary significantly from the calculated theoretical ratios. There are two main causes of this discrepancy; the first is attributed to loss of genetic homogeneity by contamination of tumor samples with normal cells from the surrounding tissue. The second discrepancy is the dependence between the fluorescence intensities of neighboring DNA fragments on the array (56) – i.e., the incorrect assumption that the measurement errors of different probes are always independent.

Some analysis algorithms require an *a priori* assumption of underlying, but unobserved, copy number states, which correspond to ‘normal’, ‘loss’, ‘gain’ etc. These blunt categories allow simple analysis of the CGH data, but are flawed, because they do not allow nuanced multifaceted effects to be determined. However, one popular algorithm, known as the hidden Markov method (HMM), instead, provides posterior distributions, which account for the dependencies in the intensity ratios of test vs. control DNA due to noise etc. and can therefore make more nuanced inferences about gains and losses in copy number. Bayesian segmentation models such as the HMM reduce noise and increase sensitivity by transforming measurements into segments of equal copy number. However, this can tend towards over-segmentation because of the assumption that the data can be modelled as a piece-wise constant (57). Despite this, the advantage of HMM is that, only normalized  $\log_2$  ratios are

required and no further tuning is necessary, which is convenient for the end-user with little or no statistical training (56).

Although Bayesian methods obviate the need for researchers to have robust statistical skills, since HMM computes posterior rather than prior distributions, it is more important to define threshold values to avoid false positive and false negative results. One threshold for deletions and duplications can be described as the mean  $\pm 4 \times \text{SD}$  (58). For an array with 3,500 loci, this cut-off level results in one false positive for every four analyses, as 99.9936% of the fragments fall within the normal range (55). For a valid diagnosis, this value of plus or minus four times SD has to be below the detection limit of an autosomal duplication. Since the  $\log_2$  of the ratio of duplication ( $\log_2 \frac{3}{2} = 0.58$ ) is closer to the normal ratio ( $\log_2 \frac{2}{2} = 0$ ) than to the ratio of a chromosomal loss ( $\log_2 \frac{1}{2} = -1$ ), this threshold can be defined as the difference between the  $\log_2$ -transformed mean intensity ratio of duplicated loci and twice the SD, i.e.,  $4 \times \text{SD} \leq ((\log_2 \frac{3}{2}) - 2 \times \text{SD})$ . In real terms, this equates to an  $\text{SD} \leq 0.096$ .

For quality control, it is imperative to report the percentage of spots with successful hybridization that provided acceptable intensity ratios (typically  $>97\%$ , signal-to-noise ratio  $>2$ ), the minimum threshold for normal clones and to confirm that the cut-off threshold for the SD of the  $\log_2$ -transformed intensity ratios did not exceed 0.096 (55).

Microarray-based assays afford many opportunities for bioinformatics and data-mining, and with this comes a pressing need for standardization. In 2001, the Functional GENomics Data Society (FGED) published the “Minimum Information About a Microarray Experiment” (MIAME) standard (59) specifying the necessary information required for unambiguous presentation and interpretation of microarray data. By 2009, most scientific journals had adopted the MIAME standard as a requirement for publishing (60) and many tailored bioinformatics tools have since been developed (e.g., TAFYYS to analyze tumor Affymetrix SNP array data (61); *equalizer* to reduce SNP bias in Affymetrix microarrays (62); Rawcopy

to improve copy number analysis with Affymetrix arrays (63); and argyle, an R Package for analyzing Illumina genotyping arrays (64).

In summary, standardization, quality control and sound biostatistics algorithms are key to publishing valid and comparable results, especially with regards to the inclusion of data in toxicologic and toxicogenomic databases, which hold a vast quantity of experimental data from different laboratories around the world.

#### ***1.4. Toxicogenomics***

The field of toxicogenomics emerged from the need to modernize traditional animal testing, which is inherently slow and expensive (65). In essence, toxicogenomics addresses the relationship between the structure/activity of the genome and the detrimental biological effects of exogenous agents, enabling a multi-endpoint analysis of previously inaccessible information about the functional activity of biochemical pathways (66). Toxicogenomics bridges genotoxicity and carcinogenicity, with methods being commercially available for high-throughput “omics” analyses to investigate how the genome and epigenome react to toxic and potentially-toxic substances (67). Consequently, genomic-scale mRNA expression (transcriptomics), cell- and tissue-wide protein expression (proteomics), as well as metabolite profiling (metabolomics) can now be more probingly evaluated with both bioinformatics and conventional toxicology (68-70).

Microarray technology in toxicogenomics allows the evaluation of toxin-modulated gene expression at the mRNA level to reveal molecular mechanisms of toxicology. With this technology, the identification of genes and gene products which may be responsible for resistance or sensitivity to toxic compounds became possible (71). Although only recently established, ‘omics’ research appears to be fulfilling its objective and revealing synergistic,



antagonistic and additive alterations in gene expression (72). The advantages of an ‘omics’ approach will most likely amalgamate the fields of cell pathology, toxicology, molecular genetics and genetic toxicology; generating a comprehensive understanding of the genetic control of normal cellular functions and of cellular responses to perturbations of those functions (66). Toxicogenomics is establishing itself as a useful tool in computational toxicology by focusing on adverse outcome pathways in chemical toxicity testing to help ensure compliance with strict regulatory frameworks for human safety testing (73). As such, the wealth of data now available in the public domain has contributed to a substantial reduction in the amount of animal testing and human clinical trials (68). Based on the analysis of gene expression using array-based toxicogenomics, it will become possible to screen for carcinogenicity and at the same time differentially analyze the possible mode(s) of action of a detrimental compound (68).

#### **1.4.1. Public Databases**

The public Comparative Toxicogenomics Database (<http://ctdbase.org/>) provides a centralized, literature-based resource that links chemical exposures via toxicogenomics to genes, proteins, genetic pathways, phenotypes and diseases to better comprehend the interaction of genes and gene products with environmental chemicals, and thus their effects on human health. This database, which includes a novel dataset on gene ontology and disease inferences, also integrates with external datasets to generate expanded networks and predict novel associations, to better understand diseases, their etiology and underlying mechanisms, upon chemical exposure (74,75). To date, the database holds close to 38 million toxicogenomics relationships. Additional to the database’s growing toxicogenomic data on various chemicals, new modules such as the “chemical-phenotype module” to better described how the effect of chemicals on phenotypes (76) or the “expanded exposure science

module” to generate predictive connections between environmental exposures and population-level health outcomes (77) are frequently introduced.

Other databases include web resources such as the toxicological data network TOXNET (<http://toxnet.nlm.nih.gov/>), the European Chemicals Agency ECHA (<https://echa.europa.eu/information-on-chemicals>), the global eChemPortal to information on chemical substances (<https://www.echemportal.org/echemportal/index.action>) or the carcinogenic potency database CPDB (<http://potency.berkeley.edu>). An overview of these and other databases can be found in a reviewed and published list of toxicology databases (78) or at <http://alttox.org/resource-center/databases/>. Alternatively, CEBS offers a comprehensive annotated database on chemical effects in biological systems (<http://www.niehs.nih.gov/research/resources/databases/cebs>) (79).

#### 1.4.2. Regulatory Guidelines

The screening of toxicants as well as the understanding of toxicity pathways based on methodologies that are sound, proven and embedded in regulatory guidelines is crucial in toxicology to establish exposure guidelines in human--health risk--assessment. Hence, the use of microarray technology marked the advent of toxicogenomics, allowing monitoring and expression-level analysis of thousands of genes on a genome--wide scale (80). In 2007, the U.S. National Research Council (NRC) released a report, "Toxicity Testing in the 21st Century: A Vision and a Strategy" that proposed a paradigm shift for testing the toxic potential of chemical substances, involving a fundamental move away from *in vivo* toxicity assessment towards *in vitro* endpoints on human cells or cell lines using high-throughput screening (81-84). Challenges with the assessment of compounds for regulatory purposes still remain, such as, off--target toxicological effects or issues of data interpretation (85) and dissemination of data or implementation of data-storage infrastructure (86). Despite these

challenges, by adopting such regulations, microarray technologies now focus more closely on human significance and have enabled a wider coverage of toxicants, a more in-depth evaluation of their effect on different life-stages, and more precise identification of sensitive subpopulations. Such regulation continues to transform toxicology testing towards the *in vitro* toxicity pathway test approach as proposed by the NRC (87,88). It has been shown that toxicogenomic tests that predict cancer or other health effects provide higher throughput; better information on the mode of action; and increased mechanistic data at a significantly lower cost compared with traditional methods (89,90). Toxicogenomics is an important adjunctive tool for traditional toxicity evaluations and there is no doubt that in concert with increased *in silico* modeling, it has led to a decrease in the use of animals for toxicity testing (91); toxicogenomics may one day replace animal testing altogether. Currently, no international regulatory guidelines on integrating toxicogenomics into human-health risk-assessment have yet been established. The need for such guidelines adapted to the challenges of the 21st century can be seen in a study by Vachon *et al.*, who found that improved reporting of methodological details and study design is key when using toxicogenomics data for human health risk-assessment (92). As more ‘omics’ data become available and the cost of implementing new experimental approaches decreases, the generation of quantitative data to assess benchmark-dose and point-of-departure dose is of increasing importance (93).

## **2. Materials**

### **2.1. DNA Labeling**

1. Nick translation buffer (10x): 500 mM Tris-HCl, pH 7.5, 100 mM MgSO<sub>4</sub>, 1 mM dithiothreitol, 500 µg/mL bovine serum albumin. For 20 mL, combine 10 mL of 1 M Tris-HCl (pH 7.5), 2 mL of 1 M MgSO<sub>4</sub>, 0.2 mL of 0.1 M dithiothreitol and 10 mg of

bovine serum albumin. Mix and adjust to final volume with pure water and freeze aliquots at -20 °C.

2. DNase I solution: 0.4 mU/μL DNase I, 40 mM Tris-HCl, pH 7.5, 6 mM MgCl<sub>2</sub>, 2 mM CaCl<sub>2</sub>. Prepare always fresh. For 15 mL of DNase buffer (10x), combine 6 mL of 1 M Tris-HCl (pH 7.5), 0.9 mL of 1 M MgCl<sub>2</sub> and 0.3 mL of 1 M CaCl<sub>2</sub>. Mix and adjust to final volume with pure water and freeze 1 mL aliquots at -20 °C. For use, dilute 10x DNase buffer, after thawing, with pure water to a 1x concentration. Then mix the DNase I enzyme (Invitrogen) with 1x DNase buffer to yield a final enzyme concentration of 0.4 mU/μL.
3. TAE buffer (50X): 2 M Tris-base, 1 M Acetic acid, 50 mM EDTA, pH 8.0.
4. 1.5% Agarose gel.
5. 0.5 μg/μL Genomic DNA.
6. 40 mM dNTPs (dATP, dCTP, dGTP, dTTP).
7. 1 mM dNTPs (dATP, dCTP, dTTP).
8. 1 mM Cy3-dCTP.
9. 1 mM Cy3-dUTP.
10. 1 mM Fluorescein-12-dATP.
11. 1 mM Fluorescein-12-dCTP.
12. 0.5 U/μL DNA polymerase I.
13. Absolute ethanol.
14. 1 kb DNA ladder
15. Loading dye
16. 0.5 μL/mL Ethidium bromide.
17. 20 mg/mL Glycogen solution.

## 2.2. Conventional CGH

All aqueous solutions have to be prepared with pure water, e.g., milliQ water.

1. 70% Formamide solution, pH 7: 70% (v/v) deionized formamide in 2x saline sodium citrate (SSC) buffer. For 100 mL, mix 70 mL deionized formamide, 10 mL 20x SSC [3M NaCl, 0.3M trisodium citrate, pH 7] and 10 mL pure water; adjust to final pH with 1 M HCl and then to final volume with pure water.
2. 50% Formamide solution, pH 7: 50% (v/v) deionized formamide in 2x SSC. For 100 mL, mix 50 mL deionized formamide, 10 mL of 20x SSC and 30 mL of pure water; adjust to final pH with 1 M HCl and then to final volume with pure water.
3. 20x SSC buffer, pH 7: 3 M NaCl, 0.3 M sodium citrate, adjust to pH 7 with 1 M NaOH. For 500 mL, dissolve 87.7 g of NaCl and 44.1 g of sodium citrate dihydrate in 400 mL of pure water; adjust to the final pH and then to final volume.
4. Master Mix: 71.4% (v/v) deionized formamide, 14.3% (w/v) Dextran sulphate, 2.86x SSC. Mix 5.5 mL of deionized Formamide, 1 mL of 20x SSC and 1 g of dextran sulphate and heat for 2--3 hours at 70 °C to completely dissolve the dextran sulphate; adjust to pH 7 with 1 M HCl and fill up to the final volume of 7 mL with pure water. Store 1 mL aliquots at -20 °C.
5. PN buffer, pH 8: 0.1 M NaH<sub>2</sub>PO<sub>4</sub> (monobasic), 0.1 M Na<sub>2</sub>HPO<sub>4</sub> (dibasic), Adjust to pH 8 with titration by adding monobasic solution (pH ~4.5) to the dibasic solution until a pH of 8 is reached, then, after measuring the reached volume, add 0.1% (v/v) IGEPAL<sup>®</sup> CA-630 (0.1 mL per 100 mL phosphate buffer).
6. PNM buffer, pH 8: 5% Non-fat dry milk powder, 0.02% sodium azide, make a suspension with PN buffer, adjust pH with 1 M NaOH while stirring, stir overnight at 37 °C, centrifuge for 10 min at 3000 x g, sterile-filter the supernatant using

0.45µm filter to avoid bacterial growth, aliquot and store working solution at 4 °C and rest of aliquots at -20 °C.

7. DNA counterstaining solution (antifade): 0.05 µg/mL 4',6-diamidino-2-phenylindole dihydrochloride (DAPI), make a solution with 2x SSC buffer and store aliquots at -20 °C.
8. Antifade solution: 0.2 M Tris-base solution, pH 8.0, 90% (w/v) glycerol, 2.33 g of 1,4-diazabicyclo[2.2.2]octane (DABCO), store aliquots at -20 °C.
9. RNase A solution: 10 mM Tris-HCl, pH 7.5, 15 mM NaCl, 10 mg/mL RNase A. Boil the stock solution for 15 min and let it slowly cool down to room temperature, store aliquots at -20 °C. Use a 1:100 dilution in 2x SSC, pH 7, as a freshly prepared working solution.
10. Pepsin stock solution: 10% Pepsin (200 mg pepsin in 2 mL pre-warmed water (37 °C); freeze 50µL aliquots at -20 °C). As a freshly prepared working solution, use 10-50 µL of pepsin stock solution per 100 mL of 0.01 M HCl, pH 2.3 (pepsin end concentration of 10-50 µg/mL).
11. 1.5 % Agarose gel.
12. Loading dye.
13. Cot-1 DNA.
14. Salmon sperm DNA.
15. Sodium acetate.
16. Magnesium chloride.
17. Formaldehyde.

## 3. Methods

### 3.1. DNA Labeling

Occasionally, the amount of isolated DNA from test cells or single cells is not enough. Then, it is recommended to amplify the DNA before DNA labeling by using degenerate oligonucleotide primed PCR [DOP-PCR (94)] or primer extension pre--amplification [PEP (95)].

#### 3.1.1. Nick Translation

1. Prerequisites: Set the water bath to 16 °C, set the heating blocks to 37 °C and 65 °C, cool down absolute ethanol at -20 °C, prepare agarose gel and buffers during the procedure (*see Note 1*).
2. Mix reactions from **Table 1** in 500µL vials, e.g., Eppendorf vials, always pipette enzymes on ice (*see Note 2*).
3. Incubate reactions at 16 °C for 90 min and then put them on ice.
4. Run 10 µL volume of each reaction on a 1.5% agarose gel (*see Section 3.1.2*).
5. Accordingly, continue incubation at 16 °C for another 15-30 min, if the fragments are still too large, or stop reaction at 65 °C for 15 min.
6. Remove excess nucleotides, by using a commercially available nucleotide removal kit (e.g., Bio-Spin<sup>®</sup> P-30 columns, Bio-Rad, or the QIAquick Nucleotide Removal Kit, QIAGEN).
7. Ethanol-precipitate the labeled DNA (*see Section 3.1.3*).

8. Keep the resuspended DNA in the fridge before use. For longer storage, store at -20 °C.

### 3.1.2. Agarose Gel Electrophoresis

1. Prepare a 1.5% normal melting point agarose gel in 1x TAE buffer.
2. For each reaction, mix 10 µL of reaction volume with 1 µL of 6x blue/orange loading dye (Promega).
3. For the size marker, mix 1 µL of a 1 kb DNA ladder (e.g., GeneRuler 1 kb DNA Ladder, ThermoFisher Scientific) with 4 µL of water and 1 µL of 6x blue/orange loading dye.
4. Fill the wells of the gel with the size marker and the reactions.
5. Use 1x TAE buffer as the electrophoresis running buffer.
6. Run the electrophoresis at 100 V, 300 mA for approximately 20 min.
7. Stain the DNA, by using 0.5 µL/mL of ethidium bromide or 0.1 µL/mL SYBR<sup>®</sup> Green I (Sigma-Aldrich), either during preparation of the gel or by submersing the gel for 10 min in the staining solution.
8. Document the agarose gel (*see Note 3*).

### 3.1.3. Ethanol Precipitation

1. To each reaction, add 1 µL of 20 mg/mL glycogen per 50 µL of volume and mix thoroughly. Glycogen will help to precipitate the DNA in an ethanol solution.
2. Add -20 °C cold ethanol (2.5 times the reaction volume).
3. Mix by inverting the vial three times.
4. Put vials for a minimum of 30 min in the freezer (better for 1 h or overnight).
5. Centrifuge at highest speed at 4 °C for 30 min using a table-top centrifuge.



6. Discard the supernatant, place the vials upside down on paper tissue.
7. Dry the pellet for 5 min using a Speed-Vac centrifuge (no heating) or at 37 °C for 20-30 min in a drying cabinet.
8. Resuspend each DNA pellet in pure water (20-80 µL) and shake at 37 °C for 30 min.

### **3.2. Chromosomal CGH**

Prerequisites: Set the water bath to 72 °C and warm up both water and Coplin jar with 70% formamide at the same time, set heating blocks to 37 °C and 76 °C, cool down 70% ethanol at -20 °C, warm up washing solutions, (a) 2x SSC, (b) 50% formamide and (c) 0.2x SSC, to 42 °C. Avoid exposure to light when working with fluorochromes.

#### **3.2.1. DNA Precipitation**

1. Mix reaction (as given in **Table 2**) in a 500µL vial, e.g., Eppendorf vial (*see Note 4*).
2. Precipitate at -20 °C overnight (alternatively at -80 °C for 30 min).
3. Centrifuge with highest speed at 4 °C for 30 min using a table--top centrifuge
4. Remove the supernatant.
5. Wash the pellet with 250 µL of 70% ethanol (-20 °C cold).
6. Centrifuge with highest speed at 4 °C for 10 min using a table--top centrifuge.
7. Remove the supernatant and place the vials upside down on tissue paper.
8. Dry the pellet for 5 min using a Speed-Vac centrifuge (no heating) or at 37 °C for 20-30 min in a drying cabinet
9. Resuspend the pellet in 3 µL of pure water. This is the “probe DNA”.

### 3.2.2. Preparation of the Hybridization Mix

1. Add 7  $\mu\text{L}$  of Master Mix (*see Note 5*) to the resuspended probe DNA (from **step 9** of **Section 3.2.1**).
2. Mix thoroughly.
3. In order to ensure a homogeneous hybridization mix, incubate at 37 °C for 30 min.

### 3.2.3. Denaturation and Pre-annealing of the Probe DNA in the Hybridization Mix

1. Denature the probe DNA at 76 °C for 7 min (do not exceed 10 min to avoid DNA fragmentation).
2. Incubate at 37 °C for 45 min to allow for pre-annealing of cot-1 DNA to repetitive DNA sequences, e.g., *alu* sequences, in the reference and test DNA (*see Note 6*).
3. In the meantime, warm up formamide solution and start denaturing metaphase slides (*see Section 3.2.5*) 15 min before the end of the pre-annealing step.

### 3.2.4. Pre-treatment of Metaphase Chromosomes with RNase A and Pepsin

*The following pre-treatment steps are optional; however, they might be crucial for optimal hybridization of the DNA probes, especially when self-made slides (96,97) are used. Pre-treatment can be done in advance and slides can be kept at -20 °C until use.*

1. Apply 150  $\mu\text{L}$  of 100  $\mu\text{g}/\text{mL}$  RNase A solution, on to the slides.
2. Cover the slides with a plastic cover slip, e.g., a piece of Parafilm.
3. Incubate at 37 °C for 60 min in a wet box.
4. Shortly dip slides in 2x SSC after removing the plastic cover slip.
5. Wash three times for 5 min in 2x SSC at room temperature.

6. Place slides in a Coplin jar filled with pepsin working solution (pre-heated to 37 °C).
7. Incubate at 37 °C for 3-10 min (see **Note 7**).
8. Rinse shortly in PBS. Wash twice for 5 min each in PBS at room temperature.
9. Incubate for 5 min in PBS substituted with 50 mM MgCl<sub>2</sub>.
10. Incubate for 4 min in 1% formaldehyde to bind the chromosomes to the glass surface and to cross-link residual proteins on the slide, this also includes the deactivation of remaining pepsin (*The incubation time might vary from 1-15 min depending on the metaphase quality*).
11. Rinse shortly in PBS, then wash for 5 min in PBS at room temperature.
12. Run slides through 70%, 90% and 100% ethanol series for 2 min each.
13. Air-dry slides (*For storage over a long period of time, keep the slides at -20 °C; then, before use, repeat steps 12 and 13*).

### **3.2.5. Denaturation of the Target DNA**

1. In a Coplin jar with 70% formamide, denature metaphase slides at 72 °C for 5 min.  
(*Caution: Temperature decreases by ca. 1 °C per slide. Denaturing of 2 to 3 slides in a 100-ml Coplin jar usually works fine; for more slides, use an additional Coplin jar. Alternatively, briefly pre-heat the slides on a heating block*).
2. Transfer the slides into a Coplin jar with 70% ethanol (-20 °C cold) and incubate for 2 min. Cold 70% ethanol is favored in this step; dehydration and the drastic temperature change helps to keep the strands of the denatured target DNA separated.
3. Incubate the slides for 2 min each in 90% and absolute ethanol (at room temperature).
4. Allow the slides to air-dry before using them in the hybridization step (*see Section 3.2.6*).

### 3.2.6. In Situ Hybridization

1. Shortly pre-warm metaphase slides on a warming block at 37 °C.
2. Apply the hybridization mix (from Step 2 in **section 3.2.3**) on to the metaphase slides at an area with a high number of metaphases.
3. Cover with a 22x22 mm<sup>2</sup> glass cover slip and seal it with rubber cement.
4. Incubate at 37 °C for 48-72 hours in an incubator.

### 3.2.7. Washing Steps

*Cautions: 1) Protect the slides from light exposure in all the subsequent steps.*

*2) Do not ever let the targeted area(s) on the slides dry out.*

1. Carefully remove the rubber cement.
2. Let the cover slip slide off in 2x SSC (this takes approx. 5 min).
3. Wash for 30 min in 50% formamide solution at 42 °C.
4. Rinse shortly in 2x SSC.
5. Wash for 5 min in 2x SSC at 42 °C.
6. Wash twice for 5 min each in 0.2X SSC at 42 °C.
7. Place slides for 5 min in PN buffer at room temperature (*See Note 8*).

### 3.2.8. DNA Counterstaining

1. Incubate slides in DNA counterstaining solution for 30-60 s at room temperature.
2. Wash shortly in pure water.
3. Place slides for 5 min each in PN buffer at room temperature.
4. Drain the buffer, apply 40 µL antifade solution and cover-slip the slides.

5. Store slides in the dark at 4 °C until evaluation.

### 3.2.9. Evaluation

1. For the evaluation of the CGH slides a fluorescence microscope, equipped with a set of suitable excitation and emission filters, CCD camera and a computer, is considered a prerequisite. In principles, it is possible to roughly analyze CGH slides manually; however, this is rather inaccurate. Analysis software is highly recommended to get an accurate quantitative and qualitative evaluation of the CGH. Companies like MetaSystems, Altlussheim, Germany offer evaluation software (<https://metasystems-international.com/us/products/solutions/cgh-telomere/>) to analyze a set of metaphases for numerical abnormalities as well as CNV as seen in loss and gains of chromosomal material. Data interpretation then becomes more precise as the results from many metaphases are integrated for each chromosome (see **Note 9**).
2. A reliable evaluation requires the assessment of a number of metaphases (usually 5 to 15) with certain characteristics. The spreading has to be optimal with an even staining and only a few or none overlapping chromosomes.
3. Besides a low unspecific background, the hybridization pattern along the chromosomes has to be even with consistent intensities.
4. The correct settings of upper and lower thresholds are therefore very important.
5. For karyotyping, the inverted black-and-white image of the counterstained metaphase is used. The chromosomes are subsequently sorted for identification according to their number, size and banding patterns by an experienced investigator. Some programs allow automatic sorting of the chromosomes, but the resulting karyotype should always be additionally checked.

6. The fluorescence intensities along the axis of the chromosomes finally result in an intensity profile. Average profiles for all 24 different human chromosomes are then calculated by using the profiles from all analyzed metaphases – at least four copies of each chromosome should be analyzed.
7. Centromeres, the p-arms of acrocentric chromosomes (13, 14, 15, 21 and 22) and heterochromatin regions, i.e., 1q12, 9q12, 16q11.2 and Yq12, are usually excluded from interpretation (*98-100,17*).

#### 4. Notes

1. Nick translation is used to incorporate fluorescently labeled deoxynucleotides into the DNA. Two enzymes, DNase I randomly nicking the DNA and bacterial DNA polymerase I, are employed in this assay. The latter enzyme possesses a 5'→3' exonuclease activity to remove 1-10 nucleotides starting from a nick in the DNA as well as a 5'→3' polymerizing gap-filling activity, elongating the DNA from the 3' hydroxyl terminus mediating nick translation along the strand. By encountering a nick on the opposite strand, a double-strand break will be generated. For DNA polymerization, this template-dependent enzyme requires deoxynucleotides (dNTP) and bivalent Mg<sup>2+</sup> ions. To enhance the efficiency of DNA labeling, the employed DNA polymerase I lacks the 3'→5' proofreading activity, which is found in the native enzyme. This assay produces a range of differently sized probes resulting in smaller fragments over time.
2. Avoid exposure to light when working with fluorochrome-labeled nucleotides. Use drawers, covers or boxes to protect the slides from direct sun light. Use yellow fluorescent lights in the lab if possible.

3. The range of the length of the labeled DNA fragments on the gel should run as a smear from 0.3–2.3 kb to ensure optimal hybridization results when used as DNA probes for CGH.
4. This DNA precipitation mix already contains all the DNAs needed for the hybridization process: both the labeled reference as well as the test DNA, human cot-1 DNA in excess (to block non-specific hybridization of the probe to non-target DNA, hence the pre-annealing step later on in the protocol) and salmon sperm DNA (which acts as a carrier DNA reducing the background by associating to non-biological sites during hybridization). Sodium acetate and ethanol are used to precipitate the DNAs.
5. The Master Mix contains 3 crucial ingredients necessary for successful hybridization of DNA probes to target DNA. Formamide destabilizes hydrogen bonds within the DNA's double helix; as a consequence, reducing the melting point of the DNA. Dextran sulphate is a polyanionic derivative of the glucan dextran acting as a crowder substance in the Master Mix. It helps to accelerate the hybridization process by reducing the access to water to nucleic acids. Monovalent cations like  $\text{Na}^+$  from the SSC buffer interact with the negatively charged phosphate backbone of the DNA, thus, increasing the affinity between the DNA probe and the target DNA.
6. This incubation is crucial as cot-1 DNA pre-anneals in this step with moderately repetitive sequences in the probe-DNA, reducing later unspecific background signals on the target DNA throughout the genome. Cot-1 DNA consist of a fraction of DNA with a re-association coefficient of  $c_o \times t = 1 \text{ mol s L}^{-1}$ , combining time of incubation and DNA concentration in this term.
7. The incubation time depends in the quality and the age of the metaphase slides. It is recommended to use different time intervals to determine the optimal incubation time

as well as different concentrations of pepsin as indicated in **Section 2.2**. Once established for a batch of slides, use the optimal concentration and time of exposure to digest superficial proteins for the rest of the slides. Avoid over-exposure to pepsin as the chromosomes can detach from the glass surface of the slide.

8. If DNA probes have been indirectly labeled with reporter molecules like biotin or digoxigenin by using correspondingly conjugated nucleotides for the mixture in Table 1, then the detection procedure must be carried out after the above washing steps. For this example, no detection procedure is necessary as the DNA probes are directly labeled with fluorochromes.
9. The resolution of an experiment is in general defined by the number of bands that can be distinguished (high resolution ranges from 550 to 850 visible bands). Hence, it is directly proportional to the length of the metaphase chromosomes. The condensation degree, which defines the length of the metaphase chromosomes can be influenced by the concentration and the duration of exposure of colcemid to the cells during the culture of PHA-stimulated peripheral blood lymphocytes. Colcemid arrests the cells in metaphase by depolymerizing the microtubules of the spindle fiber. Commercially available metaphase spreads usually show long prometaphase chromosomes, suitable for high resolution CGH.

## 5. References

1. Pinkel D, Straume T, Gray JW (1986) Cytogenetic Analysis Using Quantitative, High-Sensitivity, Fluorescence Hybridization. *P Natl Acad Sci USA* 83(9):2934-2938.
2. Ratan ZA, Zaman SB, Mehta V et al. (2017) Application of Fluorescence In Situ Hybridization (FISH) Technique for the Detection of Genetic Aberration in Medical Science. *Cureus* 9(6):e1325.



3. Heng HH, Squire J, Tsui LC (1992) High-resolution mapping of mammalian genes by in situ hybridization to free chromatin. *Proc Natl Acad Sci U S A* 89(20):9509-9513.
4. Speicher MR, Gwyn Ballard S, Ward DC (1996) Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat Genet* 12(4):368-375.
5. Schrock E, du Manoir S, Veldman T et al. (1996) Multicolor spectral karyotyping of human chromosomes. *Science* 273(5274):494-497.
6. Tanke HJ, Wiegant J, van Gijlswijk RP et al. (1999) New strategy for multi-colour fluorescence in situ hybridisation: COBRA: COmbined Binary RAtio labelling. *Eur J Hum Genet* 7(1):2-11.
7. Bolzer A, Kreth G, Solovei I et al. (2005) Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol* 3(5):e157.
8. Cremer T, Cremer C (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* 2(4):292-301.
9. Deng W, Shi X, Tjian R et al. (2015) CASFISH: CRISPR/Cas9-mediated in situ labeling of genomic loci in fixed cells. *Proc Natl Acad Sci U S A* 112(38):11870-11875.
10. Chen B, Gilbert LA, Cimini BA et al. (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155(7):1479-1491.
11. Speicher MR, Carter NP (2005) The new cytogenetics: blurring the boundaries with molecular biology. *Nat Rev Genet* 6(10):782-792.
12. Kallioniemi A, Kallioniemi OP, Sudar D et al. (1992) Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 258(5083):818-821.
13. du Manoir S, Speicher MR, Joos S et al. (1993) Detection of complete and partial chromosome gains and losses by comparative genomic in situ hybridization. *Hum Genet* 90(6):590-610.
14. Riegel M (2014) Human molecular cytogenetics: From cells to nucleotides. *Genet Mol Biol* 37(1 Suppl):194-209.
15. Kallioniemi OP, Kallioniemi A, Sudar D et al. (1993) Comparative genomic hybridization: a rapid new method for detecting and mapping DNA amplification in tumors. *Semin Cancer Biol* 4(1):41-46.
16. Zitzelsberger H, Lehmann L, Werner M et al. (1997) Comparative genomic hybridisation for the analysis of chromosomal imbalances in solid tumours and haematological malignancies. *Histochem Cell Biol* 108(4-5):403-417.
17. Piper J, Rutovitz D, Sudar D et al. (1995) Computer image analysis of comparative genomic hybridization. *Cytometry* 19(1):10-26.
18. Corso C, Parry EM (1999) The application of comparative genomic hybridization and fluorescence in situ hybridization to the characterization of genotoxicity screening tester strains AHH-1 and MCL-5. *Mutagenesis* 14(4):417-426.

19. Carlson KM, Gruber A, Liliemark E et al. (1999) Characterization of drug-resistant cell lines by comparative genomic hybridization. *Cancer Genet Cytogenet* 111(1):32-36.
20. Corso C, Parry JM (2004) Comparative genomic hybridization analysis of N-methyl-N'-nitrosoguanidine-induced rat gastrointestinal tumors discloses a cytogenetic fingerprint. *Environ Mol Mutagen* 43(1):20-27.
21. Payne J, Jones C, Lakhani S et al. (2000) Improving the reproducibility of the MCF-7 cell proliferation assay for the detection of xenoestrogens. *Sci Total Environ* 248(1):51-62.
22. Kim YM, Yang S, Xu W et al. (2008) Continuous in vitro exposure to low-dose genistein induces genomic instability in breast epithelial cells. *Cancer Genet Cytogenet* 186(2):78-84.
23. Wong N, Lai P, Pang E et al. (2000) Genomic aberrations in human hepatocellular carcinomas of differing etiologies. *Clin Cancer Res* 6(10):4000-4009.
24. Clarke PA, te Poele R, Wooster R et al. (2001) Gene expression microarray analysis in cancer biology, pharmacology, and drug development: progress and potential. *Biochem Pharmacol* 62(10):1311-1336.
25. Solinas-Toldo S, Lampel S, Stilgenbauer S et al. (1997) Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* 20(4):399-407.
26. Ylstra B, van den Ijssel P, Carvalho B et al. (2006) BAC to the future! or oligonucleotides: a perspective for micro array comparative genomic hybridization (array CGH). *Nucleic Acids Res* 34(2):445-450.
27. Brennan C, Zhang Y, Leo C et al. (2004) High-resolution global profiling of genomic alterations with long oligonucleotide microarray. *Cancer Res* 64(14):4744-4748.
28. Chan VSW, Theilade MD (2005) The use of toxicogenomic data in risk assessment: A regulatory perspective. *Clin Toxicol* 43(2):121-126.
29. Davies JJ, Wilson IM, Lam WL (2005) Array CGH technologies and their applications to cancer genomes. *Chromosome Res* 13(3):237-248.
30. Amin RP, Hamadeh HK, Bushel PR et al. (2002) Genomic interrogation of mechanism(s) underlying cellular responses to toxicants. *Toxicology* 181-182:555-563.
31. Gerhold D, Lu MQ, Xu J et al. (2001) Monitoring expression of genes involved in drug metabolism and toxicology using DNA microarrays. *Physiol Genomics* 5(4):161-170.
32. Aradhya S, Lewis R, Bonaga T et al. (2012) Exon-level array CGH in a large clinical cohort demonstrates increased sensitivity of diagnostic testing for Mendelian disorders. *Genet Med* 14(6):594-603.
33. Wang J, Zhan H, Li FY et al. (2012) Targeted array CGH as a valuable molecular diagnostic approach: experience in the diagnosis of mitochondrial and metabolic disorders. *Mol Genet Metab* 106(2):221-230.

34. Hu DG, Webb G, Hussey N (2004) Aneuploidy detection in single cells using DNA array-based comparative genomic hybridization. *Mol Hum Reprod* 10(4):283-289.
35. Cheng J, Vanneste E, Konings P et al. (2011) Single-cell copy number variation detection. *Genome Biol* 12(8):R80.
36. Fiegler H, Geigl JB, Langer S et al. (2007) High resolution array-CGH analysis of single cells. *Nucleic Acids Res* 35(3):e15.
37. Crotwell PL, Hoyme HE (2012) Advances in whole-genome genetic testing: from chromosomes to microarrays. *Curr Probl Pediatr Adolesc Health Care* 42(3):47-73.
38. Auer H, Newsom DL, Nowak NJ et al. (2007) Gene-resolution analysis of DNA copy number variation using oligonucleotide expression microarrays. *BMC Genomics* 8:111.
39. Zhao XJ, Li C, Paez JG et al. (2004) An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. *Cancer Res* 64(9):3060-3071.
40. Le Scouarnec S, Gribble SM (2012) Characterising chromosome rearrangements: recent technical advances in molecular cytogenetics. *Heredity (Edinb)* 108(1):75-85.
41. Schillert A, Ziegler A (2012) Genotype calling for the Affymetrix platform. *Methods Mol Biol* 850:513-523.
42. Hester SD, Reid L, Nowak N et al. (2009) Comparison of comparative genomic hybridization technologies across microarray platforms. *J Biomol Tech* 20(2):135-151.
43. Herzog CR, Desai D, Amin S (2006) Array CGH analysis reveals chromosomal aberrations in mouse lung adenocarcinomas induced by the human lung carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Biochem Biophys Res Commun* 341(3):856-863.
44. Medlin JF (1999) Timely toxicology. *Environ Health Perspect* 107(5):A256-258.
45. Heinloth AN, Shackelford RE, Innes CL et al. (2003) Identification of distinct and common gene expression changes after oxidative stress and gamma and ultraviolet radiation. *Mol Carcinog* 37(2):65-82.
46. Heinloth AN, Shackelford RE, Innes CL et al. (2003) ATM-dependent and -independent gene expression changes in response to oxidative stress, gamma irradiation, and UV irradiation. *Radiat Res* 160(3):273-290.
47. Brown N, Finnon R, Manning G et al. (2015) Influence of radiation quality on mouse chromosome 2 deletions in radiation-induced acute myeloid leukaemia. *Mutat Res Genet Toxicol Environ Mutagen* 793:48-54.
48. Castagnola P, Malacarne D, Scaruffi P et al. (2011) Chromosomal aberrations and aneuploidy in oral potentially malignant lesions: distinctive features for tongue. *BMC Cancer* 11:445.

49. Fujimoto J, Kadara H, Men T et al. (2010) Comparative functional genomics analysis of NNK tobacco-carcinogen induced lung adenocarcinoma development in Gprc5a-knockout mice. *PLoS One* 5(7):e11847.
50. Auerbach SS, Shah RR, Mav D et al. (2010) Predicting the hepatocarcinogenic potential of alkenylbenzene flavoring agents using toxicogenomics and machine learning. *Toxicol Appl Pharmacol* 243(3):300-314.
51. Iwahashi H, Kitagawa E, Suzuki Y et al. (2007) Evaluation of toxicity of the mycotoxin citrinin using yeast ORF DNA microarray and Oligo DNA microarray. *BMC Genomics* 8:95.
52. Huang Y, Fernandez SV, Goodwin S et al. (2007) Epithelial to mesenchymal transition in human breast epithelial cells transformed by 17beta-estradiol. *Cancer Res* 67(23):11147-11157.
53. Hong HJ, Koom WS, Koh WG (2017) Cell Microarray Technologies for High-Throughput Cell-Based Biosensors. *Sensors (Basel)* 17(6):E1293.
54. Grant GR, Manduchi E, Stoeckert CJ, Jr. (2007) Analysis and management of microarray gene expression data [chapter 19, unit 6]. *Curr Protoc Mol Biol* 77(1):1-30.
55. Vermeesch JR, Melotte C, Froyen G et al. (2005) Molecular karyotyping: array CGH quality criteria for constitutional genetic diagnosis. *J Histochem Cytochem* 53(3):413-422.
56. Guha S, Li Y, Neuberg D (2008) Bayesian Hidden Markov Modeling of Array CGH Data. *J Am Stat Assoc* 103(482):485-497.
57. Zhang L, Zhang L (2013) Use of autocorrelation scanning in DNA copy number analysis. *Bioinformatics* 29(21):2678-2682.
58. Shaw-Smith C, Redon R, Rickman L et al. (2004) Microarray based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. *J Med Genet* 41(4):241-248.
59. Brazma A, Hingamp P, Quackenbush J et al. (2001) Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* 29(4):365-371.
60. Brazma A (2009) Minimum Information About a Microarray Experiment (MIAME)--successes, failures, challenges. *ScientificWorldJournal* 9:420-423.
61. Liu Y, Li A, Feng H et al. (2015) TAFFYS: An Integrated Tool for Comprehensive Analysis of Genomic Aberrations in Tumor Samples. *PLoS One* 10(6):e0129835.
62. Quigley D (2015) Equalizer reduces SNP bias in Affymetrix microarrays. *BMC Bioinformatics* 16:238.
63. Mayrhofer M, Viklund B, Isaksson A (2016) Rawcopy: Improved copy number analysis with Affymetrix arrays. *Sci Rep* 6:36158.

64. Morgan AP (2015) argyle: An R Package for Analysis of Illumina Genotyping Arrays. *G3 (Bethesda)* 6(2):281-286.
65. Tice RR, Austin CP, Kavlock RJ et al. (2013) Improving the human hazard characterization of chemicals: a Tox21 update. *Environ Health Perspect* 121(7):756-765.
66. Aardema MJ, MacGregor JT (2002) Toxicology and genetic toxicology in the new era of "toxicogenomics": impact of "-omics" technologies. *Mutat Res* 499(1):13-25.
67. Ruden DM, Gurdziel K, Aschner M (2017) Frontiers in Toxicogenomics in the Twenty-First Century-the Grand Challenge: To Understand How the Genome and Epigenome Interact with the Toxic Environment at the Single-Cell, Whole-Organism, and Multi-Generational Level. *Front Genet* 8:173.
68. Mahadevan B, Snyder RD, Waters MD et al. (2011) Genetic toxicology in the 21st century: reflections and future directions. *Environ Mol Mutagen* 52(5):339-354.
69. OECD (2009) Series on Testing and Assessment No. 100 - Report of the Second Survey on Available Omics Tools (ENV/JM/MONO(2008)35). Available at [http://www.oecd.org/officialdocuments/displaydocument/?cote=env/jm/mono\(2008\)35&doclanguage=en](http://www.oecd.org/officialdocuments/displaydocument/?cote=env/jm/mono(2008)35&doclanguage=en). Accessed: 22 June 2018.
70. Karahalil B (2016) Overview of Systems Biology and Omics Technologies. *Curr Med Chem* 23(37):4221-4230.
71. Vrana KE, Freeman WM, Aschner M (2003) Use of microarray technologies in toxicology research. *Neurotoxicology* 24(3):321-332.
72. Song Y, Asselman J, De Schamphelaere KAC et al. (2018) Deciphering the Combined Effects of Environmental Stressors on Gene Transcription: A Conceptual Approach. *Environ Sci Technol* 52(9):5479-5489.
73. Jeong J, Choi J (2017) Use of adverse outcome pathways in chemical toxicity testing: potential advantages and limitations. *Environ Health Toxicol* 33(1):e2018002.
74. Mattingly CJ, Rosenstein MC, Davis AP et al. (2006) The comparative toxicogenomics database: a cross-species resource for building chemical-gene interaction networks. *Toxicol Sci* 92(2):587-595.
75. Davis AP, Grondin CJ, Johnson RJ et al. (2017) The Comparative Toxicogenomics Database: update 2017. *Nucleic Acids Res* 45(D1):D972-D978.
76. Davis AP, Wieggers TC, Wieggers J et al. (2018) Chemical-Induced Phenotypes at CTD Help Inform the Predisease State and Construct Adverse Outcome Pathways. *Toxicol Sci* 165(1):145-156.
77. Grondin CJ, Davis AP, Wieggers TC et al. (2018) Accessing an Expanded Exposure Science Module at the Comparative Toxicogenomics Database. *Environ Health Perspect* 126(1):014501.
78. Young RR (2002) Genetic toxicology: web resources. *Toxicology* 173(1-2):103-121.

79. Lea IA, Gong H, Paleja A et al. (2017) CEBS: a comprehensive annotated database of toxicological data. *Nucleic Acids Res* 45(D1):D964-D971.
80. Nuwaysir EF, Bittner M, Trent J et al. (1999) Microarrays and toxicology: the advent of toxicogenomics. *Mol Carcinog* 24(3):153-159.
81. Andersen ME, Krewski D (2009) Toxicity testing in the 21st century: bringing the vision to life. *Toxicol Sci* 107(2):324-330.
82. Testing CoT, Assessment of Environmental Agents NRC (2007) *Toxicity Testing in the 21st Century: A Vision and a Strategy*. The National Academies Press, Washington, D.C.
83. Andersen ME, Al-Zoughool M, Croteau M et al. (2010) The future of toxicity testing. *J Toxicol Environ Health B Crit Rev* 13(2-4):163-196.
84. Andersen ME, Krewski D (2010) The vision of toxicity testing in the 21st century: moving from discussion to action. *Toxicol Sci* 117(1):17-24.
85. Jacobs A (2009) An FDA perspective on the nonclinical use of the X-Omics technologies and the safety of new drugs. *Toxicol Lett* 186(1):32-35.
86. Hendrickx DM, Boyles RR, Kleijnans JC et al. (2014) Workshop report: Identifying opportunities for global integration of toxicogenomics databases, 26-27 June 2013, Research Triangle Park, NC, USA. *Arch Toxicol* 88(12):2323-2332.
87. Bhattacharya S, Zhang Q, Carmichael PL et al. (2011) Toxicity testing in the 21 century: defining new risk assessment approaches based on perturbation of intracellular toxicity pathways. *PLoS One* 6(6):e20887.
88. Krewski D, Westphal M, Al-Zoughool M et al. (2011) New directions in toxicity testing. *Annu Rev Public Health* 32:161-178.
89. Moffat I, Chepelev N, Labib S et al. (2015) Comparison of toxicogenomics and traditional approaches to inform mode of action and points of departure in human health risk assessment of benzo[a]pyrene in drinking water. *Crit Rev Toxicol* 45(1):1-43.
90. Chepelev NL, Moffat ID, Labib S et al. (2015) Integrating toxicogenomics into human health risk assessment: lessons learned from the benzo[a]pyrene case study. *Crit Rev Toxicol* 45(1):44-52.
91. Jean-Quartier C, Jeanquartier F, Jurisica I et al. (2018) In silico cancer research towards 3R. *BMC Cancer* 18(1):408.
92. Vachon J, Page-Lariviere F, Sirard MA et al. (2018) Availability, Quality, and Relevance of Toxicogenomics Data for Human Health Risk Assessment: A Scoping Review of the Literature on Trihalomethanes. *Toxicol Sci* 163(2):364-373.
93. Farmahin R, Williams A, Kuo B et al. (2017) Recommended approaches in the application of toxicogenomics to derive points of departure for chemical risk assessment. *Arch Toxicol* 91(5):2045-2065.

94. Telenius H, Carter NP, Bebb CE et al. (1992) Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. *Genomics* 13(3):718-725.
95. Zhang L, Cui X, Schmitt K et al. (1992) Whole genome amplification from a single cell: implications for genetic analysis. *Proc Natl Acad Sci U S A* 89(13):5847-5851.
96. Huber R, Kulka U, Lorch T et al. (2001) Technical report: application of the Metafer2 fluorescence scanning system for the analysis of radiation-induced chromosome aberrations measured by FISH-chromosome painting. *Mutat Res* 492(1-2):51-57.
97. Bangs CD, Donlon TA (2005) Metaphase chromosome preparation from cultured peripheral blood cells. *Curr Protoc Hum Genet* Chapter 4:Unit 4 1.
98. du Manoir S, Kallioniemi OP, Lichter P et al. (1995) Hardware and software requirements for quantitative analysis of comparative genomic hybridization. *Cytometry* 19(1):4-9.
99. du Manoir S, Schrock E, Bentz M et al. (1995) Quantitative analysis of comparative genomic hybridization. *Cytometry* 19(1):27-41.
100. Lundsteen C, Maahr J, Christensen B et al. (1995) Image analysis in comparative genomic hybridization. *Cytometry* 19(1):42-50.

## Figure Captions

### Figure 1:

Chromosomal CGH – Representation of gain (on p-arm) and loss (terminal; on q-arm) of chromosomal material in the test genome in relation to the reference / control genome.

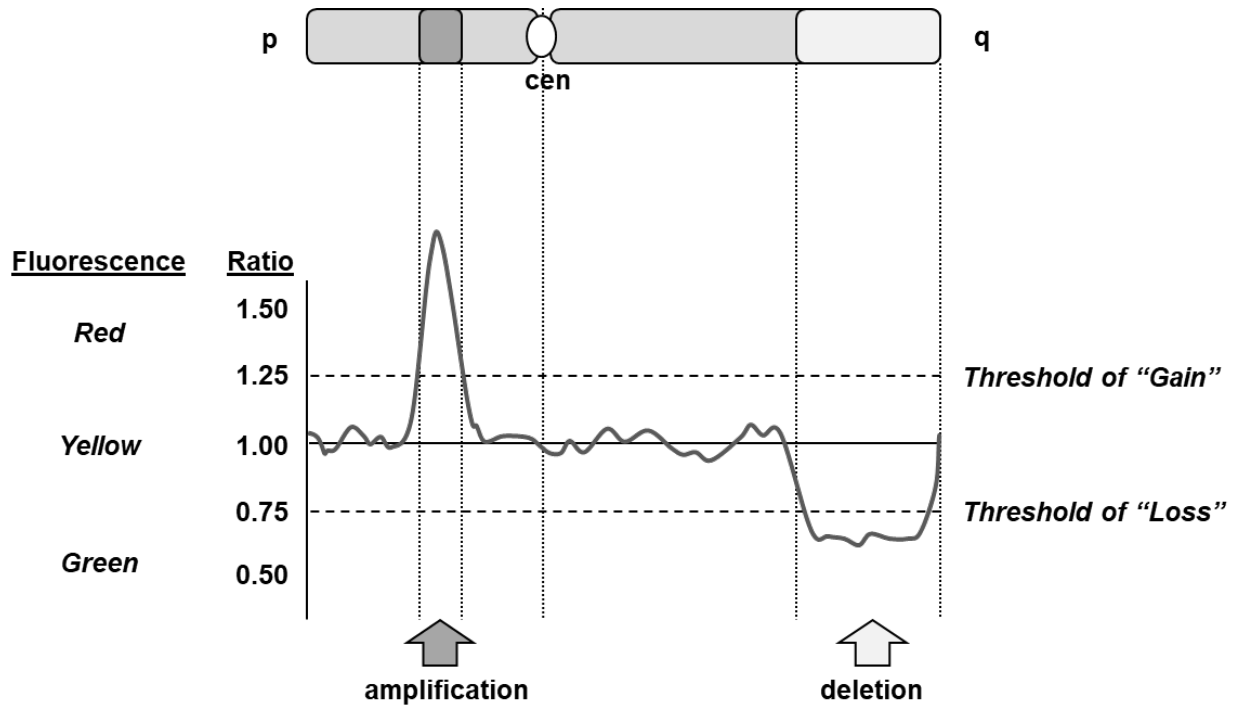
Genomic DNA is isolated from the test and control tissue, labeled with green and red fluorochromes, respectively. After hybridization, the fluorescence ratio of  $1.0 = \frac{2}{2}$  indicates two copies of homologue sequence throughout the chromosome – for both the control and test genomic DNA (resulting in the additive color, yellow). For instance, the loss of one homologue leads to a ratio of  $0.5 = \frac{1}{2}$ , below the threshold of “loss” (the lower dashed line in the figure). The latter then only shows the fluorescence of the labeled control DNA (green). Amplification of chromosomal material results in a fluorescence ratio of  $1.5 = \frac{3}{2}$  (e.g., partially trisomic resulting in a red-shift). Note that the ratios shown in this figure have not undergone base-2 logarithmic transformation.

### Figure 2:

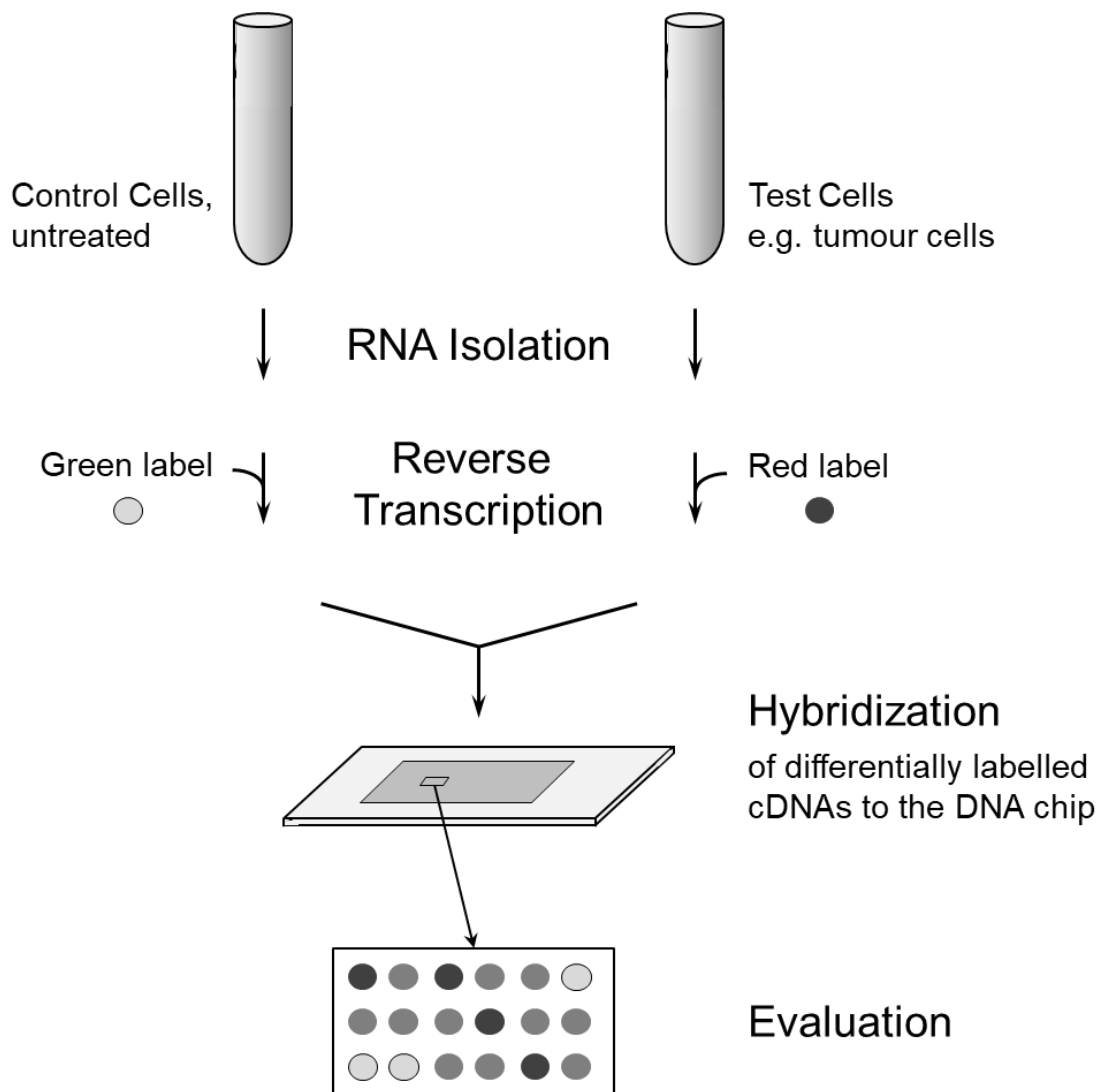
Array CGH. Fluorescently labeled cDNAs generated from RNA of both control and test cells are hybridized onto a DNA chip (e.g., oligonucleotide microarray). Subsequent evaluation of the resulting dot matrix on the chip produces the gene expression profile of the studied cells.



**Figure 1:**



**Figure 2:**



## **Table Captions**

### **Table 1:**

Exemplar mixtures for labeling normal and test genomic DNA by using Nick translation. The test DNA can be for instance from a tumor.

### **Table 2:**

Exemplar mixture of various probe DNA for CGH

**Table 1:**

Reagents	Reference DNA [ $\mu$ L]	Test DNA [ $\mu$ L]
Water (pure, autoclaved)	53.4	53.4
Genomic DNA (0.5 $\mu$ g/ $\mu$ L)	12	12
Nick-Translation Buffer (10x)	10	10
dATP (40 mM)	–	2
dGTP (40 mM)	2	2
dTTP (40 mM)	2	–
dATP (1 mM)	0.8	–
dCTP (1 mM)	0.8	0.8
dTTP (1 mM)	–	0.8
Cy3-dCTP (1 mM) [PerkinElmer]	–	4
Cy3-dUTP (1 mM) [PerkinElmer]	–	4
Fluorescein-12-dATP (1 mM) [PerkinElmer]	4	–
Fluorescein-12-dCTP (1 mM) [PerkinElmer]	4	–
DNase I Solution (5-10 $\mu$ L depending on DNA)	5	5
DNA Polymerase I (0.5 U/ $\mu$ L)	6	6
Total Volume	100	100

**Table 2:**

Reagents	[ $\mu$ L]
Reference DNA, fluorescein-labeled	20
Test DNA, rhodamine-labeled	20
Human cot-1 DNA (1 mg/mL) [Invitrogen]	40
Salmon sperm DNA (10 mg/mL) [Invitrogen]	1
Sodium acetate (3 M, pH 5.0); 1/10 volume	8.1
Ethanol, absolute, -20 °C; 2.5x volume	202.5
Total Volume	291.6