Université de Montréal

Méthode d'étude moléculaire des tumeurs rares pédiatriques

Methods of molecular analysis to study rare paediatric tumours

Par
Papp Eniko

Département de pathologie et biologie cellulaire
Faculté des études supérieures

Mémoire présenté à la Faculté des études supérieures
En vue de l'obtention du grade de Maîtrise en science (M.Sc)
En Pathologie et Biologie Cellulaire
Option cancérologie

Août 2006

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé :
Méthode d'étude moléculaire des tumeurs rares pédiatriques

Methods of molecular analysis to study rare paediatric tumours

Présenté par :
Papp Eniko

A été évalué par un jury composé des personnes suivantes:

Dr Daniel Sinnett
.....
Président rapporteur

Dr Jean-Christophe Fournet
.....
Directeur de recherche

Dr Raouf Fetni
.....
Membre du jury

SOMMAIRE :

Après le décès d'origine accidentelle, les cancers représentent la principale cause de mortalité chez les enfants et adolescents de moins de 20 ans. Seule les 5 cancers les plus fréquents dans cette tranche d'âge sont relativement bien caractérisés sur le plan moléculaire. Tout reste à faire pour les 200 restant. Malheureusement, leur étude reste difficile en raison de nombreuses limitations. Premièrement, l'accès aux échantillons tumoraux frais et congelés est limité du fait même de la rareté de ces tumeurs. Deuxièmement, l'obtention de mitose, indispensable pour l'étude des aberrations chromosomiques, est difficile. Troisièmement, l'étude de l'ADN à partir de prélèvements d'archive est réduite en raison de la dégradation de l'ADN au sein de ces prélèvements inclus en paraffine. Ainsi, dans l'état actuel de nos connaissances, la quantité d'information disponible sur la génétique de ces tumeurs reste faible. Afin d'améliorer l'étude de ces tumeurs et pour acquérir de nouvelles informations sur les voies oncogénétiques moléculaires, nous nous proposons de combiner des méthodes classiques d'analyse (FISH interphasique, allélotypage des tumeurs à l'aide de marqueurs microsatellites) avec une nouvelle technologique de génotypage tumoral a haut débit. En combinant ces méthodes, nous contournons le problème de mitose et la dégradation en utilisant la méthode de FISH interphasique et nous permettons l'étude a haute débit même avec peu de matériel en utilisant la plateforme Illumina. En premier lieu, nous avons mis au point une technique de génotypage en utilisant la plateforme Infinium d'Illumina et la puce Sentrix Human-1 SNP. Pour ce faire, nous avons utilisé 25 échantillons de néphroblastome (tumeur de Wilms') comme référence, ce modèle tumoral ayant déjà fait l'objet d'un allélotypage avec une carte de délétion du chromosome 11p. Une fois notre technique validé, nous avons étudié deux tumeurs orphelines à titre de modèles, afin de comparer des techniques génétiques classiques et une méthode de génotypage tumoral de haute performance: la myofibromatose infantile comme premier modèle et la tumeur musculaire lisse associée au virus rapporteur comme deuxième modèle. Notre but est de perfectionner les approches diagnostiques et thérapeutiques des tumeurs pédiatriques rares et, à terme, de permettre une amélioration de la survie des patients. Mots clés : Cancer pédiatrique rare, tumeur de Wilms', myofibromatose infantile, tumeur musculaire lisse associée au virus d'Epstein-Barr, hybridation in situ en fluorescence, allèlotypage, perte d'hétérozygotie, plateforme Illumina.

SUMMARY:

After prenatal complications and accidental deaths, cancer is the most common disease-related cause of death among children and adolescents younger than 20 years of age. Even if the 5 most common paediatric cancers are well characterised, the molecular pathways involved in oncogenesis still remain unknown in more than 200 paediatric tumours. The reason of this major lack in knowledge is due to limitations when studying rare paediatric and orphan tumours. First, there is a limited access to fresh and frozen cancer tissue samples due to the rarity of these tumours. Second, there is limited amount of mitosis achieved which is however necessary in the study of chromosomal aberration. Finally, the study of the DNA is limited due to the degradation of the DNA in paraffin embedded tissues. Due to these reasons, the study of rare paediatric cancer is challenging, limiting our knowledge about their genetic information. To overcome these problems, we propose that combining low-throughput methods (interphasic FISH which does not require mitosis, molecular allelotyping studies using microsatellite markers) with a new high- throughput tumoral genotyping method could improve characterization of samples and increase our knowledge of the molecular oncogenesis. By combining these methods, we are able to study tumor samples contained in paraffin where there is no mitosis by using iFISH. Furthermore, we are able to study small amount of tumoral samples by using high throughput genotyping Illumina platform. To test this hypothesis, we used 25 diagnosed nephroblastoma (Wilms') tumour samples. Chromosome 11p allelotyping was carried out with these samples. With the detailed deletion map as a reference, we validated the Illumina platform and the Sentrix Human-1 SNP BeadChip array. Afterwards, we used two orphan tumours as models to analyse classical genetic and the high-throughput genotyping techniques: Infantile Myofibromatosis as our first model and Intestinal Epstein-Barr virus-associated smooth muscle as our second model. Our goal is to improve diagnostic and therapeutic approaches and, ultimately, enhance individual clinical outcomes.

Key words: Rare paediatric cancer, Wilms' tumour, Infantile Myofibromatosis, Intestinal Epstein-Barr virus-associated smooth muscle, fluorescent in situ hybridization, allelotyping, loss of heterozygosity, Illumina platform, Sentrix Human-1 SNP BeadChip array.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

ABREVIATIONS

% : percent

µ : micro

°C: degree celcius

ALK: anaplastic lymphoma kinase

ASPE : allele-specific primer extension

ATBF1:AT-binding transcription factor 1

BAC : bacterial artificial chromosome

bp : base pair

caR : dinucleotide CA repeat region

CEPH : centre d'étude du polymorphism humaine

CGH : comparative genome hybridization

CHU : centre hospitalier universitaire

del : deletion

der: derivé

DM : double minute

DNA : Deoxyribonucleic acid

EBV : Epstein-Barr virus

EBV-SMT: EBV-smooth muscle tumour

FANCA: Fanconi anaemia complementation group A

FISH : fluorescence *in situ* hybridization

FUS : http://www.gene.ucl.ac.uk/nomenclature/data/get_data.php?hgnc_id=4010 fusion (involved in t(12;16) in malignant liposarcoma)

g : gram

gDNA : genomic DNA

HD-CGH: high definition CGH

HPLC: high performance liquid chromatography

HSR: homogeneously staining regions

IFM : Infantile Myofibromatosis

ICB: illumina chromosome browser

IGF2: insulin-like growth factor 2

IGV: illumina genome viewer

Kb : 1 thousand base pairs

l : liter

Log: logarithm

LOH: loss of heterozygosity

LOI:	loss of imprinting
M:	mitosis
Mb :	1 million base pairs
M-FISH:	multiplex-FISH
ml:	mililiter
MMP2:	matrix metallopeptidase 2
mol :	mole
myc:	myelocytic leukemia
MYCN:	myc myelocytomatosis viral related oncogene, neuroblastoma derived
N:	number
NCBI :	national center for biotechnology information
NCI:	national cancer institute
ng:	nanogram
NQO1:	NAD(P)H dehydrogenase quinine
ONC:	oncogene
p:	short arm of the chromosome
PAC:	P1-derived artificial chromosome
PCR:	polymerase chain reaction
pmol:	picomole
POU6F2:	the POU domain, class 6, transcription factor 2
PTCH:	patched homolog drosophila
q:	long arm of the chromosome
RASSF1:	Ras association (RalGDS/AF-6) domain family 1
RBL2:	retinoblastoma like-2
RFLP:	restriction fragment length polymorphisms
SKY:	spectral karyotyping
SNP:	single nucleotide polymorphism
t:	translocation
TAQ :	thermophilic DNA polymerase
TSG:	tumour suppressor gene
U :	unit
UPD:	uniparental disomy
WGG :	whole genome genotyping
WT:	wild type
YAC:	yeast artificial chromosome

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor, Dr. Jean-Christophe Fournet. I could not have imagined having a better clinical advisor and mentor, and without his boldness (who knew Illumina platform existed!), common-sense, knowledge and perceptiveness, I would have never completed my master's project.

I express my gratitude to Dr. Daniel Sinnett for his time, constructive advice and to always have considered me as part of the team, well adopted student that is! Also, I will be forever thankful for Dr. Labuda's generosity and the support I had from his laboratory team.

A special thank you for Dr. Raouf Fetni who introduced me the cytogenetic world and to have shared his knowledge about the clinical side of research.

Dr Luc Oligny, thank you for always being proud of your students! I would also like to thank all the rest of the academic and support staff of the department of pathology and the Hemato-Oncology department, particularly those who have put up with my drifting a long way away from my original title. Thank you Vania for always being there and for the support you have gave me.

Finally thank you for all my friends and family, wherever they are, particularly my Mum and Dad; and, most importantly of all, to Dorothée, Anne-Laure and Karine for being a support throughout this journey. A special thank you for Alexandre Montpetit for his constant help and patience during data analysis.

Peace and Solidarity,

Eniko

CHAPTER I

- Introduction -

During the last several decades, intensive research efforts have increased our understanding of carcinogenesis and have identified that cancer development is caused by a multi-step process. Tumors grow through a process of clonal expansion driven by mutations generating a group of cells genetically identical to a single common ancestor. Clonal expansion is a property obtained after the cancer cell accumulated mutations leading to a survival advantage. The first somatic or constitutional mutation in an oncogene or tumour-suppressor gene causes a clonal expansion, thus initiating the neoplastic process subsequently followed by additional mutations causing further rounds of clonal expansion. The genetic abnormalities generated by these mutations contribute to cancer pathogenesis including self-sufficiency, angiogenesis and metastasis (Figure 1) [1, 2].

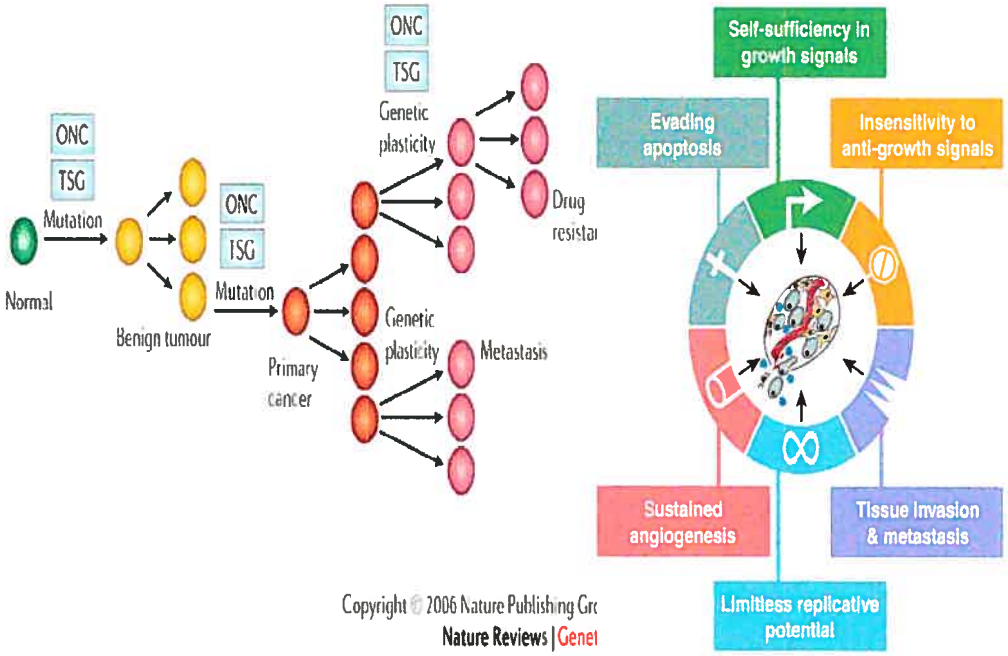


Figure 1: *The multi-steps of cancer. Cancer arises through a series of mutations, involving among other oncogenes (ONC) and -suppressor genes (TSG). Each mutation leads to the selective overgrowth of a monoclonal population of cells, as well as to significant properties (invasiveness, metastasis, drug resistance, etc.) [1, 2].*

Advances in cancer research have allowed us to gain knowledge about the molecular level of different adult cancer types including but not limited to breast, lung, colon and ovarian cancer. However, paediatric cancers are less well characterized and remain an important cause of death among children. After accidental deaths, cancer is the most common disease-related cause of death among children and adolescents younger than 20 years of age (The national cancer institute of Canada, 2006). Leukemias, lymphomas, central nervous system, neuroblastomas and Wilms' tumours are the most frequently diagnosed cancer categories and are relatively well characterized. However, more than 200 tumours observed in children still need to be characterized in order to provide a better diagnosis and treatment to children affected by cancer.

We have considered a paediatric tumour to be rare if its incidence is lower than Wilms' (0.8 cases per 100 000 persons; 7% of paediatric solid tumors). We defined that an orphan is a tumoral type with less than 20 PubMed references in Cytogenetics or molecular biology. Examples of such tumours are listed in Table I and II.

The goal of this study was to compare various types of techniques in order to gain knowledge about the molecular and cytogenetic information on rare and orphan paediatric cancers. We wanted to have an overall idea of the amount of information that can be obtained when studying rare and orphan tumours with limited genetic information.

***Table I**: List of common (in green) and rare paediatric (in red) cancers with the number of PubMed references for each type. The columns represent the number of reference found in pubmed when gene, chromosome, LOH, cytogenetics or CGH was typed in the pubmed search.*

	Tumor	Global	Gene	Chromosome	LOH	Cytogenetics	CGH
Common paediatric tumors	neuroblastoma	23 737	4 732	1 201	178	80	47
	osteosarcoma	18 922	2 638	719	27	74	32
	astrocytoma	18 255	2 587	780	131	35	40
	paraganglioma	16 414	870	220	40	4	14
	Wilms tumour	8 297	1 780	999	76	49	11
Cutoff							
Rare paediatric tumors	rhabdomyosarcoma	8 995	850	430	24	58	14
	hepatoblastoma	1 895	454	112	8	9	8
	Ewing sarcoma	1 162	255	158	1	21	6

***Table II**. Some examples of paediatric orphan tumours with less than 20 PubMed references in cytogenetics/oncogenetics (shown in red). The columns represent the number of reference found in pubmed when gene, chromosome, LOH, cytogenetics or CGH was typed in the pubmed search.*

Tumor	Global	Gene	Chromosome	LOH	Cytogenetics	CGH	Allelotyping
desmoid tumour	1029	88	34	0	1	0	0
clear cell sarcoma of kidney	201	24	15	0	2	1	0
gastrointestinal stromal tumour	1381	221	46	4	0	5	0
Hemangio-endothelioma	2750	38	19	1	0	1	0
insulinoma	4234	734	76	14	2	1	2
lipofibromatosis	7	0	0	0	0	0	0
pancreatic acinar cell carcinoma	210	21	5	0	0	1	1
solid pseudopapillary tumour of the pancreas	163	10	2	0	1	1	0

Mutations generating the genetic abnormalities contributing to cancer pathogenesis can be classified into two main groups (illustrated in figure 2): constitutional mutations gained via germline cell mutations; somatic mutations arising in one cell, from which will ultimately originate the monoclonal proliferation. Thus, the methods employed to collect data in oncogenetics must study the constitutional genome and the tumoral genome. In this study, we have tried to use as much of these techniques to better characterise rare and orphan tumours.

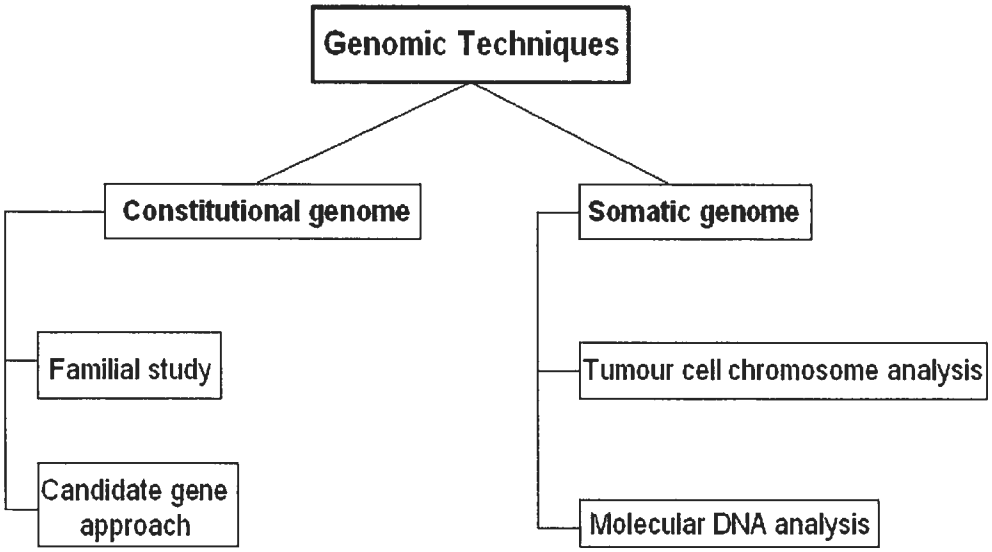


***Figure 2**: Diagram illustrating various genomic techniques used to study both the constitutional and somatic tumour genome.*

1. Study of the constitutional genome

1.1. Familial Studies

DNA sequences can be analyzed and compared by powerful genetics tools including genome wide linkage scans.

Linkage analysis exploits the phenomenon of recombination to localise disease-genes relative to the known position of genetic markers [3]. The segregation of each marker is compared to the segregation pattern of the disease phenotype within an affected family and measured by a logarithm of the odds (LOD) ratio. This calculates the likelihood that the marker is segregating with the disease gene if there is a lack of recombination. Although linkage analysis can pinpoint the region containing the cancer predisposition gene to a domain much smaller than a chromosomal band, identification of the predisposition gene ultimately requires positional cloning approaches and detailed mutational analyses. Linkage analysis is a powerful tool in the hunt for cancer genes, however, it requires a large number of families to carry out the analysis, thus limiting the study of hereditary cases. Nevertheless, linkage analysis helped to clone many hereditary cancer genes over the years, including tumour suppressor genes [4-6].

Tumour suppressor genes protect the cells from deregulated growth and division by controlling cell cycle progression or by driving damaged cells into apoptosis [7]. Knudson's epidemiological studies on retinoblastoma led to the proposal of the "two-hit" model [8]. He prposed that retinoblastoma is a cancer caused by two mutational events causing the inactivation of a tumour suppressor gene. In the dominantly inherited form, one mutation is inherited via the germinal cells and the second occurs in somatic cells. In the nonhereditary form, both mutations occur in somatic cells [8]. Patients with a hereditary mutation of a tumour suppressor gene are at a much greater risk of developing tumors than the general population because the probability of acquiring a single somatic mutation is exponentially greater than the probability of acquiring two such mutations [9].

Later on, in 1997, Kinzler and Vogelstein refined this model by introducing the knowledge of caretakers and gatekeepers. Unlike the mutation-based definition of tumour suppressor genes, the division is based on function and have added complexity and power to the original hypothesis. Gatekeepers are genes that directly regulate the initiation of tumors by inhibiting growth of malignant cells or promoting death [10]. In contrast, inactivation of a caretaker gene does not promote tumour initiation directly. Rather, inactivation leads to genetic instabilities which result in increased mutation of all genes, including gatekeepers [10].

1.2. Constitutional mutations screening in candidate gene

Another way to discover potential cancer causing genes is by candidate gene mutation screening. After the identification of a putative candidate gene, PCR amplifications of the coding DNA are performed in order to detect mutations. A variety of techniques can be used to perform the mutation screening including but not limited to: sequencing, denaturing high performance liquid chromatography (HPLC), and protein truncation test (PTT). Identifying mutations in several unrelated affected individuals strongly suggests that the correct candidate gene has been chosen. After identification of mutations, several subsequent steps are required to further investigate the role of the mutated gene in cell lines or in animal models. Identification of novel genes driving cells to become cancerous can lead to improved diagnosis and patient counselling. (http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hmg.section.1959).

2. Study of somatic genome

2.1. Analysis of cell chromosomes (tumoral cytogenetics)

2.1 a) Chromosomes in tumoral cells

Double helix DNA is tightly compacted in a cell since it is wrapped around histone proteins. The nucleosome consists of 1.65 turns of DNA wrapped around the histone octamer complex which is further compacted into chromatin. During cell division, the chromatin strands become more and more condensed causing transcription to stop. This compact form makes the individual chromosomes visible, and they form the classic four arm structure, a pair of sister chromatids attach to each other at the centromere. The shorter arms are called *p arms* and the longer arms are called *q arms* (figure 2A). Humans have 23 pairs of chromosomes, which makes the *diploid* number 46. The diploid number is the number of chromosomes of a normal cell. Chromosomes can be visualized by a technique knows as a karyotype.

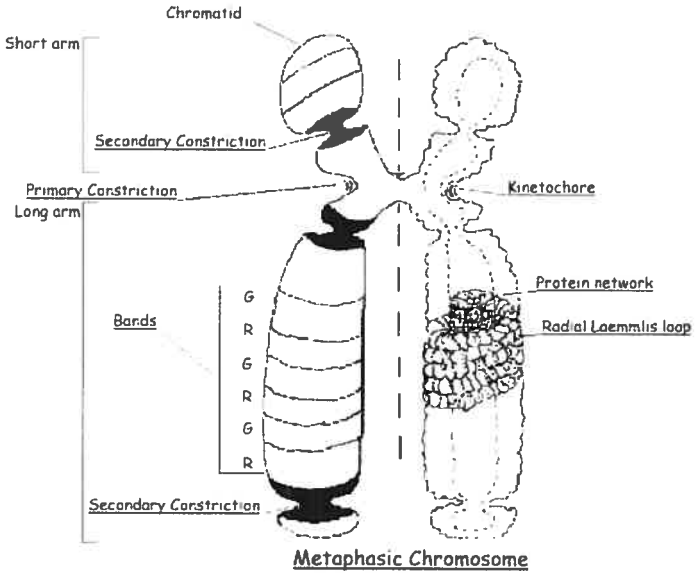


Figure 2A. Metaphasic chromosome structure showing the *p arms* and the *q arms*.

A karyotype, generally done in but not limited to GTG banding, is a technique that allows geneticists to visualize metaphase chromosomes under a microscope. The discovery of the correct number of chromosomes in humans was achieved in 1956 [11]. This finding revolutionalized the world of genetics and allowed the association of well-recognized syndromes, such as Down syndrome, Turner syndrome and Klinefelter syndrome with their specific chromosomal anomalies [12]. In cancer, one of the exciting findings was the identification in 1960 of a minute chromosome, later named the Philadelphia chromosome, which was regularly found in the peripheral blood of patients with chronic myeloid leukaemia [13].

Karyotypes can be performed on virtually any population of rapidly dividing cells either grown in tissue culture or extracted from tumours. However, one of the disadvantages of this technique is the difficulty to grow the tissue cultures. Karyotype analysis can also be performed with cells but requires up to two weeks to obtain a sufficient amount of cells for analysis. Even with these limitations, karyotypic abnormalities have been described in more than 10 000 human neoplasms analyzed by means of chromosome banding [14]. These aberrations are of two kinds: structural chromosome abnormalities (translocation, deletion, isochromosome, amplification) and numerical chromosome abnormalities (aneuploidies for example trisomie).

i Translocation

A chromosomal translocation is a rearrangement in which part of a chromosome is detached by breakage and subsequently joined to a non-homologous chromosome [15]. Chromosome translocation can occur by various mechanisms and require a double-strand break at the chromosome level. Double-strand breaks can arise upon replication across a nick, and hence, all factors that lead to a nick might ultimately yield some double-strand breaks [16]. In addition, factors that cause a substantial amount of direct double-strand break formation also may cause translocations. These include ionizing radiation, oxidative agents, and enzymes (such as type II topoisomerases and recombinases) [16].

Even if all genes are present in normal dosages, translocations or even inversions can alter the phenotype because of subtle position effects. The structural changes may involve an equal exchange of material between two chromosome regions (balanced chromosome translocation) or may be non-reciprocal, such that portions of the genome are lost or gained (unbalanced chromosome translocation) [17].

Balanced rearrangements, that have been characterized molecularly act by deregulating a gene in one of the breakpoints or by creating a fusion gene [18]. Balanced, reciprocal translocations are common in haematological malignancies, while in solid tumours (sarcomas or carcinomas) virtually all rearrangements are unbalanced resulting in loss/gain of chromosome parts [19-21]. A large number of lymphoid malignancies are characterized by the activation of a silent proto-oncogene through its relocation at the vicinity of an active regulatory element [22]. Proto-oncogenes are key players in the control of normal cell growth and proliferation [1].

Several strategies were developed to identify oncogenes, representing mutated forms of proto-oncogenes, in human tumours. Methods include DNA transfection techniques and mapping of breakpoints in non-random chromosomal rearrangements (*The genetic basis of human cancer, Chapter 10*; illustrated in Table IV). The mechanisms that cause these translocations remain poorly understood. Illegitimate V(D)J recombination, class switch recombination,

homologous recombination, non-homologous end-joining and genome fragile sites all have potential roles in the production of non-random chromosomal translocations [15]. More than 100 oncogenes have been identified and associated with some form of cancer.

Table III: *Mapping of some breakpoints in non-random chromosomal rearrangements allowed oncogene identification (Table from* The genetic basis of human cancer, Chapter 11: Oncogenes, Morag Park, p. 211*)*

Affected Gene	Translocation	Disease	Protein Type
Gene fusion: c-ABL (9q34) BCR (22q11)	t(9;22)(q34;q11)	Chronic myelogenous leukemia	Tyrosine kinase activated by BCR
Oncogene juxtaposed with IG Loci: BCL-2	t(14;18)(q32;q21)	Follicular lymphoma	Inner mitochondrial membrane
Oncogene juxtaposed with TCR loci: c-myc	t(8;14)(q24;q11)	Acute T-cell leukemia	HLH domain
Gene fusion in sarcomas: WT1, EWS	t(11;22)(p13;q12)	Desmoplastic small round cell	Wilms' gene
Oncogenes juxtaposed with other loci: BTGI / MYC	t(8;12)(q24;q22)	B-cell chronic lymphocytic leukemia	MYC-HLH domain

ii Deletion

A chromosomal deletion can occur on any chromosome and can vary by size. Consequently, the outcome of the deletion depends on the genes contained in the lost region. The loss of chromosomal regions containing a tumour suppressor gene is a key event in the evolution of epithelial and mesenchymal tumors and it is further discussed later on in the text [23].

iii Isochromosome

Isochromosome formation results when one arm of a chromosome is lost and the remaining arm is duplicated, resulting in a chromosome only consisting of either two short arms or of two long arms. An isochromosome has identical genetic information in both arms. Isochromosomes are observed in 10% of cancer cytogenetic examinations (illustrated in table III).

***Table IV**: Examples of isochromosome-associatedfor example in solid tumours or leukemias, table from: www.humpath.com*

acute myeloid leukemia	i(11q)	i(17q)	i(21q)		
chronic myeloid leukemia	i(9q)	i(17q)	i(22q)		
chronic myeloproliferative disorders	i(17q)				
myelodysplastic syndromes	i(X)(q13)	i(17q)	i(21q)		
acute lymphoblastic leukemia	i(7q)	i(9q)	i(17q)		
chronic lymphoproliferative disorders	i(1q)	i(7q)	i(8q		
Hodgkin disease	i(1q)	i(6p)	i(9p)	i(17q)	i(21q)
non-Hodgkin lymphoma	i(1q)	i(6p)	i(17q)		
adenocarcinoma	i(1q)	i(8q)	i(17q)		
transitional cell carcinoma	i(5p)	i(8q)	i(11q)		
Wilms tumor	i(1q)	i(7q)	i(17q)		
germ cell neoplasms	i(1q)	i(12p)	i(17q)		
sarcoma	i(1p)	i(1q)	i(6p)	i(17q)	
mesothelioma	i(5p)	i(6p)	i(7p)	i(21q)	
malignant neurogenic neoplasms	i(1q)	i(6p)	i(17q)		
retinoblastoma	i(1q)	i(6p)	i(17q)		
malignant melanoma	a	i(1q)	i(6p)	i(8q)	

iv Amplification

Another method transforming a proto-oncogene to an oncogene is by amplification of restricted regions of the genome. Amplification represents one of the major molecular pathways by which gene expression is constitutively enhanced above the level of physiologically normal variation [24]. Increase of the gene dosage by DNA amplification is a common genetic mechanism for upregulating gene expression in tumorigenesis. Amplification values usually are above five copies, often can reach 500 or more gene copies [24]. Neuroblastoma patients' prognosis can be determined by the copy number of the MYCN (myc myelocytomatosis viral related oncogene, neuroblastoma derived) oncogene [25]. Another example of an oncogene amplification is the EGFR (epidermal growth factor receptor) amplified in 34% of glioblastomas and constitutes a potential target for therapy [26].

Amplification of an oncogene can be caused by an unscheduled replication during cell division or if part of the DNA is excised following loop formation (Figure 3) [24]. Chromosomal regions which are amplified can be visualized by double minute chromosomes (DMs) or homogeneously staining regions. DMs are small, paired, usually spherical chromatin bodies consisting of genes amplified in an extrachromosomal location [27]. In a later step, they may integrate into chromosomes to generate intrachromosomally amplified structures known as homogeneously staining regions (HSRs). On the other hand, amplifications known as tandem amplification can occur in situ, not extrachromosomally, and can be tandem repeats of varying sizes. DMs and HSRs can be found in virtually any type of solid human tumors.

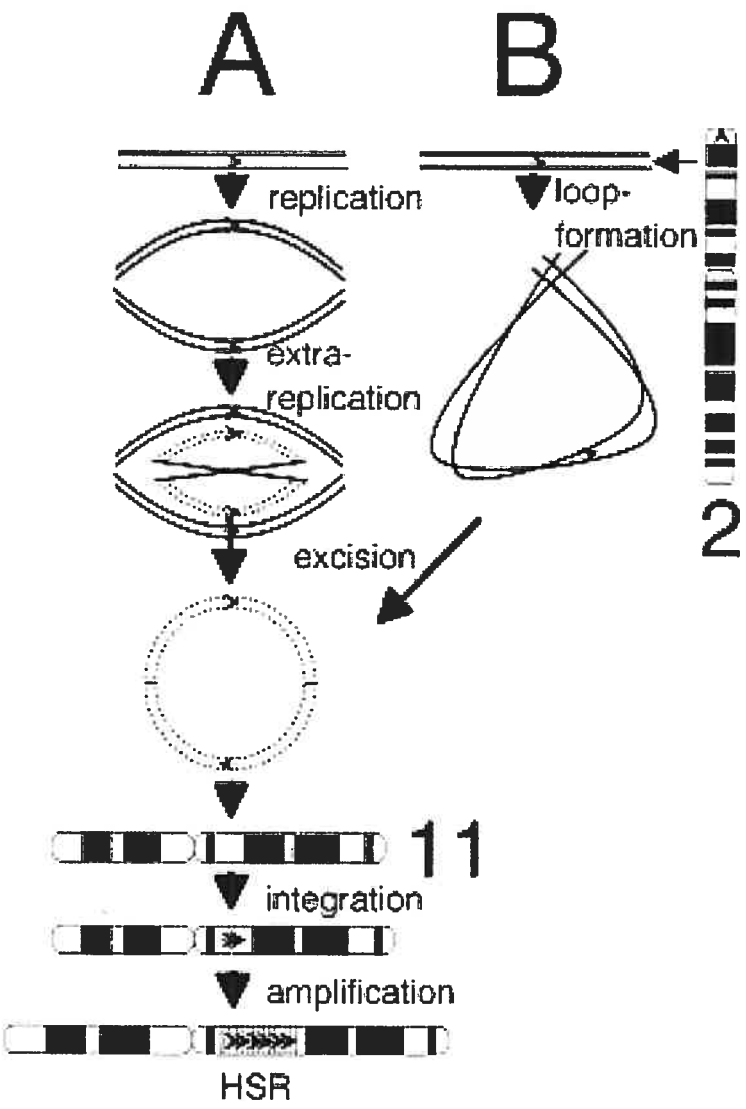


***Figure 3**: DNA amplification. Unscheduled replication during cell division (A) or partial DNA excision following loop formation (B) can results but is not limited in DNA amplification. In normal cells the MYCN gene is localized as a single copy on chromosome 2, whereas cells with amplification have up to several hundred copies, generally located in an HSR on another chromosome (in this case chromosome 11)* [24].

v Trisomies

The altered transcript levels in cancer genome can be caused by altered gene copy number changes such as a gain of an individual chromosome. Previous results from laboratories have implicated a role of chromosomal trisomies in various cancers. Hepatoblastoma, the most frequent malignant liver tumour in children, is associated with recurring trisomies of chromosomes 2, 8, and 20 [28].

2.1 b) Cytogenetic tools in cancer genetics

i. Molecular Cytogenetic

Cytogenetic approaches are designed to detect aberrations and rearrangements under direct examination of chromosomes and chromosomal targets [29]. G-banding, fluorescence in situ hybridization (FISH), spectral karyotyping (SKY) and comparative genomic hybridization (CGH) are the commonly used methods.

ii. G-banding

G-banding evaluates stained metaphase chromosome spreads to identify rearrangements and gains or losses of chromosome bands. The major advantages of G-banding consist of the stability and high resolution of the Giemsa staining; it is also easy to use, highly reproducible, and inexpensive [30]. However, there are major disadvantages to this technique. It requires the analysis of metaphase spreads which are not always possible to obtain due to cell-culture failure or poor chromosome morphology. Also, it has a limited resolution of one band (about 2000~4000 kb in size) and cannot be used to detect very small deletions implicated in many microdeletion syndromes [31]. Finally, it cannot be used to study interphase cells. In recent years, new molecular cytogenetic technologies have rapidly developed and revolutionized the diagnostic applications of chromosome analysis.

iii. FISH

Fluorescence *in situ* hybridization (FISH) is a powerful molecular cytogenetic technique which allows rapid detection of aneuploidies, translocations, on metaphase spreads and interphase cells. FISH technology uses probes that are specific for the sequence of interest and evaluates alterations at the specific locus on a cell-by-cell basis. FISH analysis provides a more accurate assessment of the entire population of cells in a given sample because both interphase and metaphase cells can be assessed. Also, it has advantages over G-banding in terms of speed and cell-scoring ability [32]. However, the numbers of commercially available probes are limited and they do not cover the whole genome. With the help of the completion of the human genome, it is now possible to use bacterial artificial chromosome (BAC) probes to surpass the lack of commercially available probes. DNA sequences spanning the human genome have been cloned and inserted into different vectors including BACs. BACs are particularly useful tools because they can divide indefinitely in culture, and can be used to package relatively long lengths of DNA sequence [33, 34]. Labelled BACs, usually by nick-translation, can be used to detect chromosomal abnormalities virtually throughout the whole genome.

Advances in FISH technology involve the use of spectral karyotyping (SKY) and multiplex-fish (M-FISH). SKY uses 24 different probe sets to virtually paint each metaphase chromosome with a different color. This technique involves the simultaneous excitation of multiple fluorochromes and the use of an interferometer to determine the profile at each pixel. Multiplex-FISH is similar to SKY except that a fluorescence microscope is used for the image analysis [13]. Multicolour FISH analysis can contribute to the diagnosis of patients with apparently normal karyotypes displaying minute abnormalities not detectable with low resolution methods [31].

iv. CGH array

The preparation of high-quality metaphase spreads, especially from solid tumours, is often difficult. To overcome this problem, comparative genomic hybridization (CGH) was developed using extracted tumoral genomic DNA. This is a powerful technique which allows the detection of segmental DNA copy number changes. Differentially labelled tumour DNA and control normal DNA are co-hybridized to a metaphase chromosome spread, producing an average fluorescence ratio profile at approximately 20 Mb resolution [13]. This resolution limitation can be improved by the use of array CGH.

In array CGH, chromosomal targets are replaced by arrays consisting of well-defined genomic clones such as Bacterial Artificial chromosome (BAC), P1-dervided artificial chromosome (PAC), both grown in bacteria or Yeast Artificial Chromosome (YAC) grown in yeast. They are spotted onto a microscopic slide-glass using robotic devices. Since the clones spotted on slide-glass contain sequence information directly connected with the genome database, we can easily obtain particular biological aspects of genes mapped within the regions involved in a copy number aberration. This facilitates the identification of genes responsible for cancer as well as unknown genetic diseases. Several companies now offer CGH platforms. For example, NimbleGen's human whole-genome array CGH platform and Agilent technologies high-definition CGH (HD-CGH) microarrays permit researchers to design their own CGH microarrays to target specific "hot spots" in the genome [35].

The principles of the CGH array are explained in Figure 4a) and 4b) and consist of the following steps:

(A) BAC clones are selected from a physical map of the genome.

(B) DNA samples are extracted from selected BAC clones and their identity is confirmed by DNA fingerprinting or sequence analysis.

(C) A multi-step amplification process generates sufficient material from each clone for array spotting. Each clone is spotted in replicate onto a solid support.

(D) Reference DNA and test DNA are differentially labeled with cyanine 3 and cyanine 5 respectively.

(E) The two labeled products are combined and hybridized onto the spotted slide.

(F) Images from hybridized slides are obtained by scanning in two channels.

(G) Signal intensity ratios from individual spots can be displayed as a simple plot.

(H). or by using more complex software that can display copy number alterations throughout the whole genome

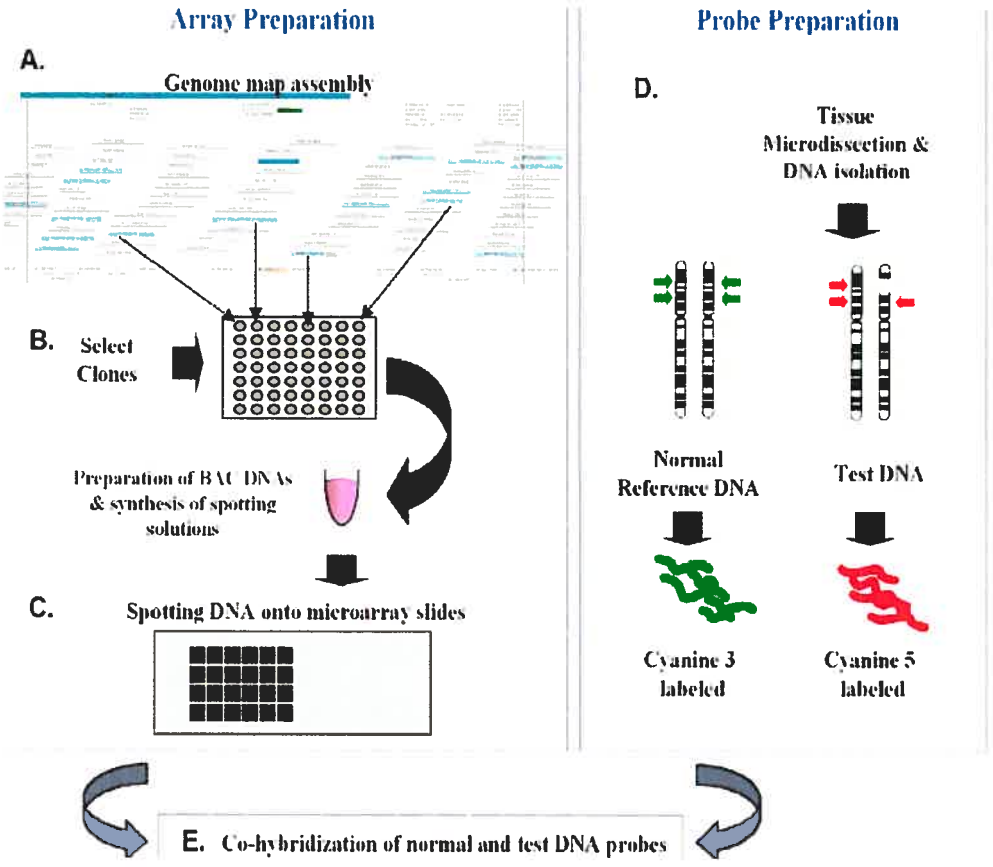


***Figure 4a**: Principle of array CGH, part 1 (Image and text form [29])*

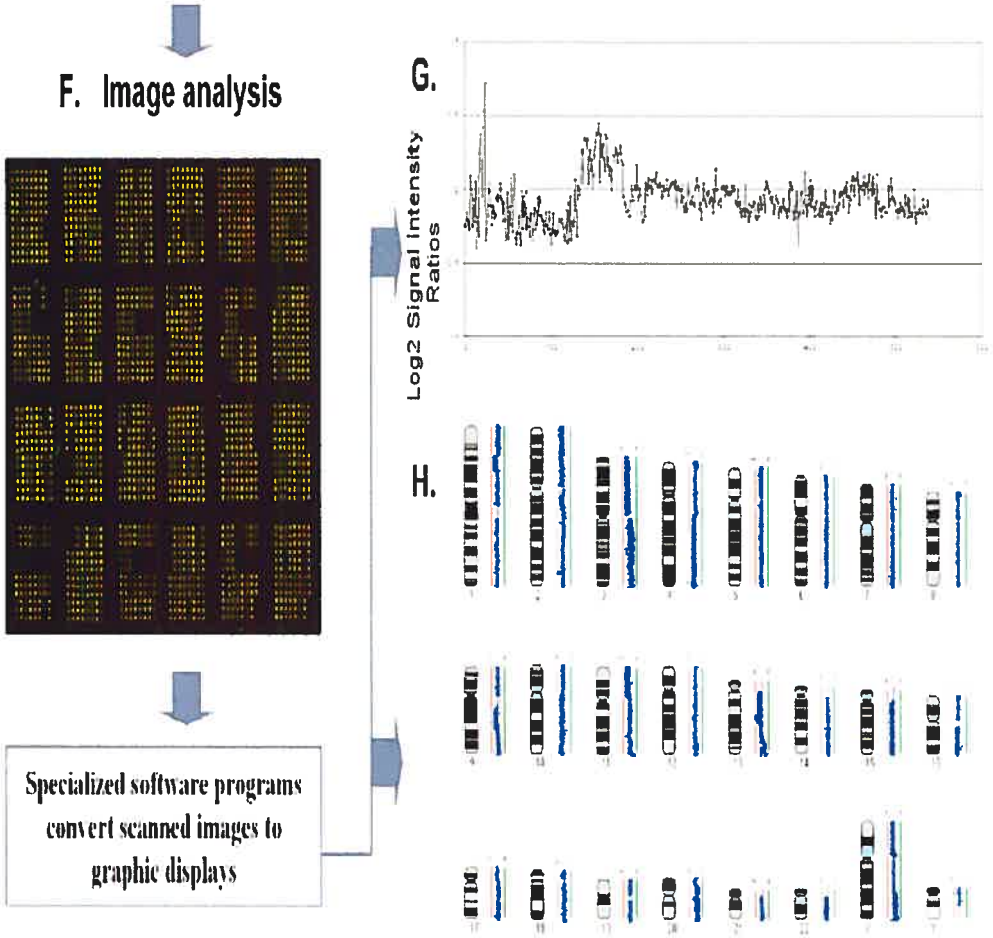


***Figure 4b**: Principle of array CGH, part 2 (Image and text form [29]).*

One of the main advantages of CGH is that it does not require an a priori knowledge of chromosome imbalance and can be used as a powerful screening tool in the field of genetics. CGH array does have some limitations: it can not detect smaller deletions than its resolution permits, balanced chromosomal translocations, inversions and whole genome copy number (ie. the loss of one allele followed by reduplication of the remaining allele is not detected [13]).

As cytogenetic methods are widely used in both clinics and in research laboratories, profile databases have been assembled for public access. The main public online databases are presented in Table V.

***Table V**: Online cytogenetic resources*

www.wiley.com/legacy/products/subject/life/mitelman	mitelman's catalogue of chromosome aberrations in cancer
www.ncbi.nlm.nih.gov/sky/skyweb.cgi	NCBI-SKY/M-FISH & CGH Database
http://amba.charite.de/cgh/	Charité - Comparative Genomic Hybridization
www.progenetix.net/	Progenetix - Online CGH Database
www.helsinki.fi/cmg	Laboratory of Cytomolecular Genetics (CMG)
http://cgap.nci.nih.gov/	NCI-CGAP (cancer genome anatomy project) The tumour Gene database

3. Studies of tumoral genome (somatic anomalies)

3.1. Molecular DNA analysis

Genomic DNA obtained from tumoral tissues is an essential asset to study genetic variations causing oncogenesis. An important event occurring in the tumoral genome is chromosome deletion involving a suppressor gene. Unlike oncogenes, for most tumour suppressor genes, both alleles need to be inactivated. TSG being recessive, cells that contain one normal and one mutated gene are referred to as heterozygous and still behave normally. However, silencing the remaining normal suppressor gene predisposes the cell to develop into a tumour following a "loss of heterozygosity" (LOH). In the case illustrated in Figure 5, the allele illustrated by the microsatellite marker caR2 and caR3 is considered heterozygous, thus informative. However, if a second hit occurs (in this case a loss of chromosomal region illustrated in yellow) the remaining tumour suppressor gene is inactivated and the cell can no longer sustain normal cell proliferation. When amplifying the microsatellite marker at this region, a LOH can be observed on the gel upon analysis of the tumoral tissue.

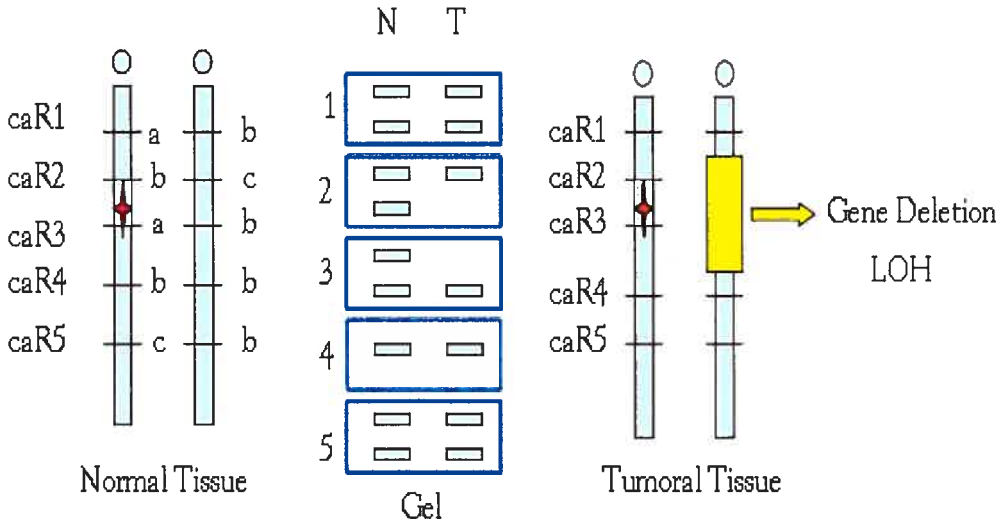


***Figure 5**: LOH causing TSG inactivation.*

A loss of DNA can occur with unbalanced chromosomal translocations, chromosomal deletions, and mitotic nondisjunctions [36, 37] (illustrated in Figure 6).

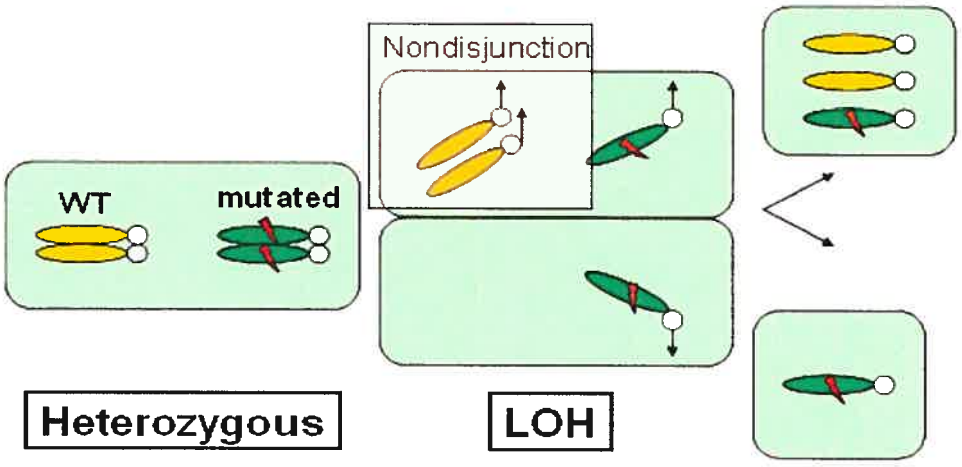


Figure 6: *Chromosomal nondisjunction. One daughter cell can inherit only one chromosome, bearing the mutant allele, while the other inherits three chromosomes (the one bearing the mutant allele as well as two normal chromosomes). In this example, LOH reflects a true reduction to hemizygosity.*

However, a hemizygous deletion leading to a copy number reduction is not the only event causing the inactivation of a tumour suppressor gene. A variety of different genetic events underlie LOH, including point mutation, mitotic nondisjunction followed by replication of the remaining chromosomes, mitotic recombination (illustrated in Figure 7) and gene conversion [38]. These LOH mechanisms do not lead to DNA copy number changes and are therefore referred to as copy-neutral events [39].

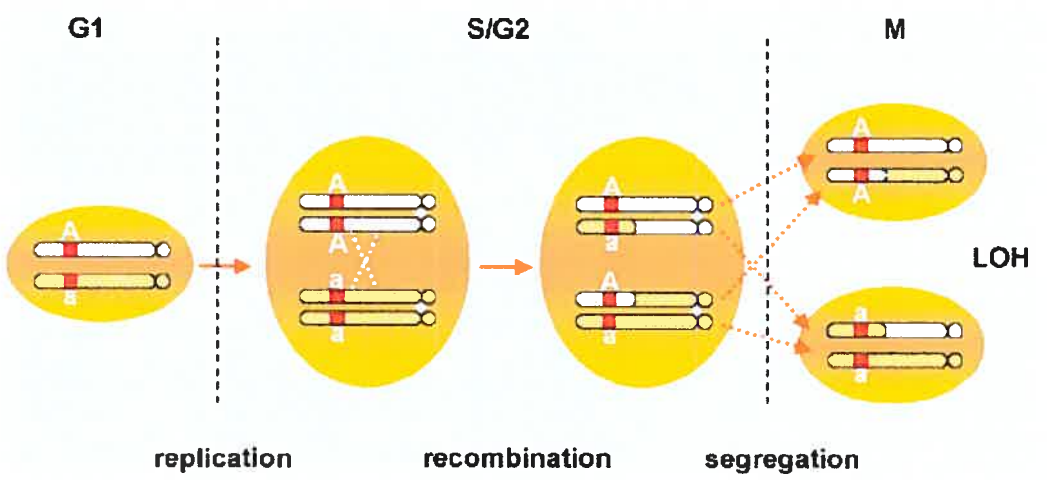


***Figure** 7: Mitotic recombination. Recombinational exchange between two sister chromatids of homologous chromosomes during late S- or G2-phase of the cell cycle causes mitotic recombination. When resolution of the recombinational intermediate results in a cross-over event, daughter cells may, following mitosis (M), show loss-of-heterozygosity (LOH) without a net reduction in chromosomal copy number [38].*

Thus, the use of the term allelic imbalance to describe LOH results may be technically more accurate. The most reliable method to characterize allelic imbalances should have the ability on one hand to provide a locus-specific genotyping and on the other hand to quantify accurately the copy number of each allele.

To date, many articles published studies on LOH without necessarily identifying definite genes causing cancer. By combining this enormous amount of data into a LOH database, including both positive and negative results would help defining the minimal region of chromosomal loss and the subsequent identification of putative tumour suppressor genes [40].

There are three major techniques used in LOH studies:

a) Southern blotting analysis of restriction fragment length polymorphism,
b) Microsatellite Marker Amplification,
c) High-resolution single-nucleotide polymorphism array.

3.1 a) Southern blotting analysis of restriction fragment length polymorphism

Loss of tumour suppressor gene function in the progression from normal to cancerous cells can be detected by assessing the loss of heterozygosity. These studies have traditionally been performed using Southern blotting-based analysis of restriction fragment length polymorphisms (RFLP) [41, 42]. This method differentiates an individual's parental alleles at a given locus and allows an assessment of LOH by simple comparison of the allelic patterns detected in matched pairs of normal and DNA samples. The study of LOH by Southern blot is limited by the low heterozygosity rates of RFLP and by the requirement for relatively large amounts of homogeneous high molecular weight DNA. It is difficult to study small endoscopic biopsy specimens or microdissected surgical resections by using southern blotting-based analysis of RFLP. Moreover, paraffin-embedded archival pathology samples are not amenable to Southern blot analysis because of extensive DNA degradation [43]. For these reasons, more efficient and less time-consuming methods are used to study LOH.

3.1 b) Microsatellite marker amplification

Microsatellites are among the most variable types of DNA sequence in the genome. Among dinucleotides, $(CA)_n$ repeats are most frequent, followed by $(AT)_n$, $(GA)_n$ and $(GC)_n$ [44]. Genetic variation at many microsatellite loci is characterized by high heterozygosity and the presence of multiple alleles. Thus, microsatellites are a great tool for loss of heterozygosity analysis. PCR amplification of samples is done by using the polymorphic repeat marker and is compared to the matched normal DNA. When LOH analysis is extended to multiple chromosomal arms, a distinct allelotype is generated [45].

Fluorescence-based DNA sequencing technologies facilitated the detection of loss of heterozygosity (LOH) by replacing the use of radionucleotide-based detection with automated generated data. This new system enables automated size determination, linear quantitation of alleles, and computerized discrimination of true alleles versus stutter bands [46].

Laser capture microdissection can now permit the obtention of highly purified samples, thus allowing a better analysis of LOH [47, 48]. However, the resolution for whole genome scanning is limited to 5 cM with commercially available sets of primers, and the process for whole genome analysis is long and tedious and requires a matching normal sample [49]. Microsatellite markers generally require individual amplification reactions or at best only a limited multiplex assay. An average of 120 microsatellites has been used to determine the allelotype of multiple different human neoplasms in a series of studies before 1995. The highest density allelotyping published before the year 2000 consisted of approximately 280 polymorphic markers [50]. Recently, a more efficient method of genome wide allelotyping by using single nucleotide polymorphism array has been developed.

3.1 c) High-resolution single-nucleotide polymorphism array

High density single-nucleotide polymorphic allele arrays permit the generation of genome-wide loss of heterozygosity maps [50-53]. Single nucleotide polymorphisms (SNPs) are the most common genetic variation in the human genome and can be used to search for germline genetic contributions to disease. The density, distribution, and allele specificity of SNPs makes them attractive for high-resolution analyses of LOH and copy number alterations in cancer genomes [54-59]. In 2004, we thought it might be interesting to use this high throughput method to study rare paediatric cancers. At the time, only a few studies were published on cancer genetics using the Affymetrix SNP array [53, 60, 61]. These studies showed that most affected LOH regions were consistent with those observed in previous LOH studies using microsatellite marker amplification, lending validity to both the methods and their results.

However, LOH studies do not clearly define a cancer gene; many candidate genes in a region of LOH are needed to be tested for mutation, amplification and promoter methylation before assigning a specific gene to a specific cancer. Methods complementary to DNA analysis, such as mRNA and protein expression microarray, will also provide important corroborative evidence [40]. These approaches can be limited when genes that do not follow the classical definition of tumour suppressor genes are studied. Indeed, many tumour suppressor genes do not fit the classical definition of tumour suppressor genes and the Knudson two-hit hypothesis (illustrated in Figure 8) [62].

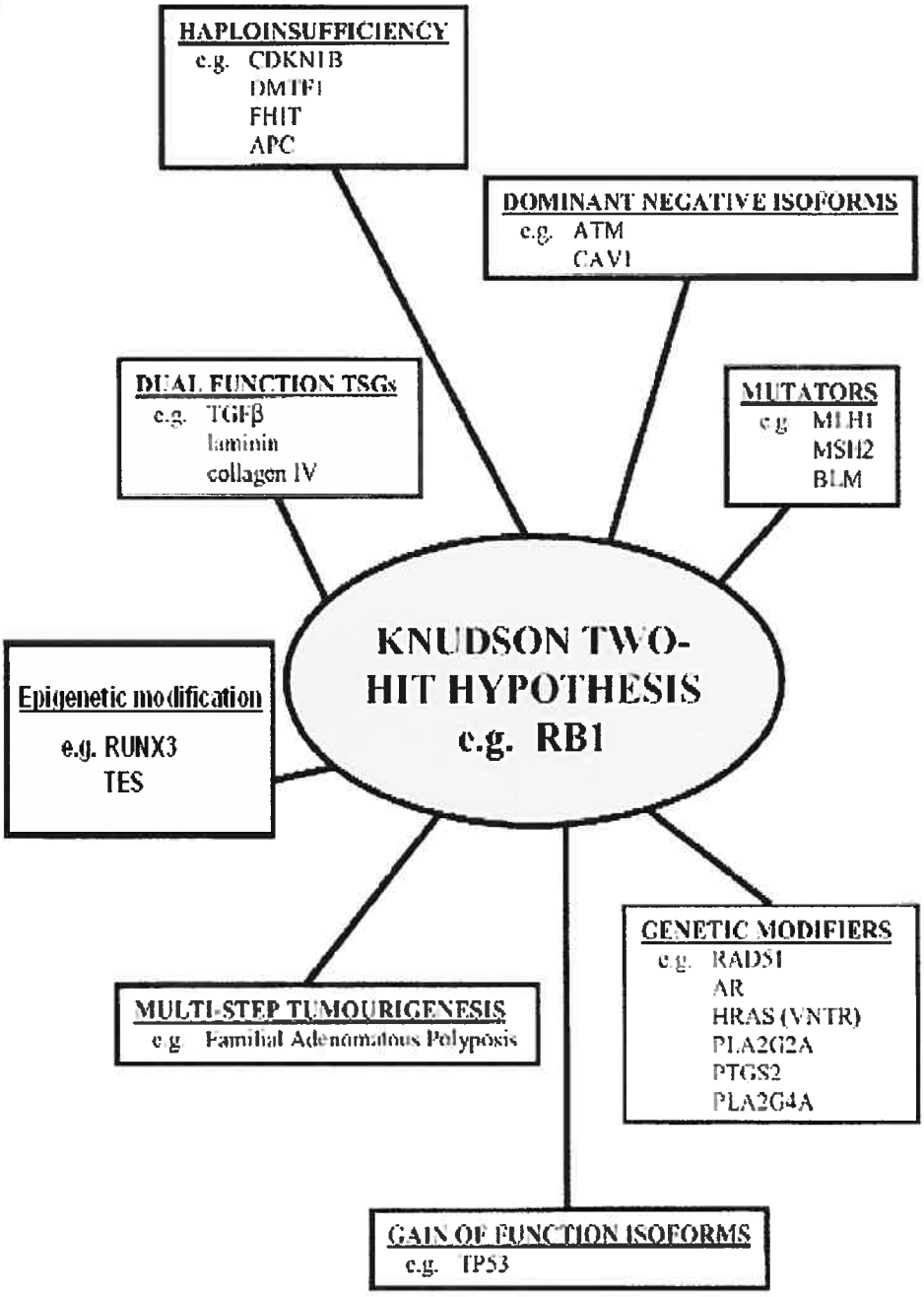


Figure 8: *Alternative mechanisms causing the inactivation of tumour suppressor genes. The illustrated mechanisms represent alternatives to the Knudson two-hit hypothesis [62].*

These genes deviate from the classic tumour suppressor gene definition in the sense that they can arise from:

(i) Monoallelic disruption (mainly explained by haploinsufficiency where a reduction of 50% in the level of gene function is sufficient to generate an abnormal cellular phenotype),
(ii) Multiple gene interaction (eg. the Familial Adenomatous Polyposis model),
(iii) Epigenetic inactivation in one or both alleles.

Because of the heterogeneous nature of the alterations observed, a comprehensive approach is required to fully characterize the molecular patterns of tumour suppressor gene inactivation. Another method, not previously discussed, has been developed to analyse genomic imprinting, also known as epigenetic inactivation.

3.1 d) Genomic imprinting and methods of detection

Certain cells of the embryo are omnipotent; their genes are all potentially active and available for transcription. During development, some genes need to be activated, whereas others are inactivated to allow differentiation [63]. However, cancer cells are able to reactivate the once-silenced genes allowing a growth benefit or to inactivate genes that hinder growth. Cancer cells essentially use the same machinery and cascades as do normal cells to benefit their growth. The epigenetic mechanisms contributing to the development of cancer include global DNA hypomethylation, hypermethylation and hypomethylation of specific genes, chromatin alterations and loss of imprinting [64].

Genomic imprinting is an epigenetic modification of a specific parental allele [64]. These modifications lead to a differential expression of the two alleles of the gene in somatic cells of the offspring. Many mammalian imprinted genes regulate cell growth, differentiation, and apoptosis. It has also been shown that genomic imprinting plays a major role in tumorigenesis. Loss of imprinting (LOI) refers to the activation of the normally silenced allele or the reverse, the silencing of a normally active allele.

Analysis of mice carrying uniparental disomies or duplications, by microarray screening and stringent bioinformatics allows a large-scale tissue-specific screening for imprinted genes [65].

CHAPTER II. HYPOTHESIS

The study of rare paediatric and orphan tumours has many limitations: they limited access to fresh or frozen cancer tissue samples, absence of mitosis to study chromosomal aberration and degradation of DNA in paraffin embedded tissue. For all to these reasons, the study of these tumours is extremely difficult limiting our knowledge about their genetic information. To overcome these problems, we hypothesize that combining low-throughput methods (iFISH, microsatellite analysis) with a new high-throughput tumoral genotyping method could improve characterization of tumour samples and increase our knowledge of molecular oncogenetic pathways.

CHAPTER III. OBJECTIVES OF THE PROJECT

The purpose of the present study was to develop new experimental methods to study rare paediatric cancers. To do so, we examined both well characterized tumours and tumours with limited molecular knowledge to determine which method is best suited to improve information gathering. Our goal is to improve diagnostic and therapeutic approaches and, ultimately, enhance individual clinical outcome.

The specific objectives were:

1. Use previously diagnosed samples of nephroblastoma (Wilms') with known chromosomal deletions to validate the Illumina Sentrix Human-1 SNP BeadChip array.

2. Use two orphan tumors as models to analyse classical genetic and high-throughput techniques:

    A. First model: Infantile Myofibromatosis,
    B. Second model: Intestinal Epstein-Barr virus-associated smooth muscle tumour

CHAPTER IV. MOLECULAR CHARACTERIZATION OF RARE PAEDIATRIC TUMOURS

A. SNP array validation

B. Infantile Myofibromatosis

C. Intestinal Epstein-Barr virus-associated smooth muscle

A. Single nucleotide polymorphism array to test LOH in Wilms' tumours

1. Rational for Wilms' selection

We chose Wilms' to test the Illumina platform because it is the most common abdominal tumour in children and it permits us to have a good amount of samples to be tested. Also, its cytogenetic and molecular data are well known permitting us to recognize known chromosomal aberrations in our Wilms' samples and allowing us to validate the SNP platform before using it to investigate our rare paediatric tumors.

Sporadic development of Wilms' tumours is associated with chromosomal deletions at both bands 11p13 and 11p15. In 45% of sporadic cases, sub-band 11p15.5 LOH is present with a selective loss of the maternal allele and a duplication of the paternal allele [36]. Other observations of LOH at 1p, 4q, 7p, 14q, 16q in sporadic Wilms' tumours identified additional regions of the genome that may harbour important genes in development [36].

With these known chromosomal aberrations in mind, we used Wilms' samples to test the Sentrix Human-1 SNP BeadChip array. First, we identified LOH in our 25 Wilms' samples (DNA was extracted for 25 patients, normal gDNA was extracted from unaffected kidney and tumoral gDNA was extracted from the tumour) by amplifying microsatellite markers on chromosome 11. Once the LOH regions were mapped, we compared the results of microsatellite amplification with the SNP array data in 6 patients, with or without LOH on chromosome 11.

2. Sample selection and DNA extraction

25 Wilms' tumours were collected from patients who underwent surgery at the CHU Ste-Justine. In each case, the diagnosis was confirmed by the Pathology Department at CHU Ste-Justine. DNA was extracted with standard phenol/chloroform purification.

3. Loss of heterozygosity (LOH) analyses by $(CA)_n$ microsatellites PCR amplification

a. Methods

Matched normal and tumoral tissues were amplified with microsatellite markers (ordered from Invitrogen and detailed in figure 9) for chromosome covering regions 11p15 to 11p13. In this approach, the forward unlabeled microsatellite primer is synthesized with an M13 forward primer sequence on the 5'-end (5'-**CAC GAC GTT GTA AAA CGA C**-3'). An M13 IRDye 800 (LI-COR part # 829-05565) was included in the PCR reaction. The M13 dye is added to the PCR product during the first few cycles of amplification thus labelling the PCR product. This method requires no purification prior to gel analysis.

PCR amplification was carried out in a 10 µl reaction volume containing: 25 ng of genomic DNA template, 1X standard buffer with 1U Platinium TAQ (Invitrogen), 0.2 µM of forward and reverse primer, 0.2 mM dNTP, and 1.0 pmol M13 FWD (-29)/IRDye 800. PCR cycles were adapted for each microsatellite, including 10-minute denaturing step at 95°C , 26-30 cycles of 15-second denaturation at 95°C, 30-second annealing step at 55°C, 1-minute elongation step at 72°C and 5-minute final extension step at 72°C. The PCR products were mixed with bromphenol blue containing loading buffer (LI-COR part # 830-05629), separated by electrophoresis on 6.5% polyacrylamide gel and detected by laser fluorescence using a LICOR automated gene sequencer (LI-COR).

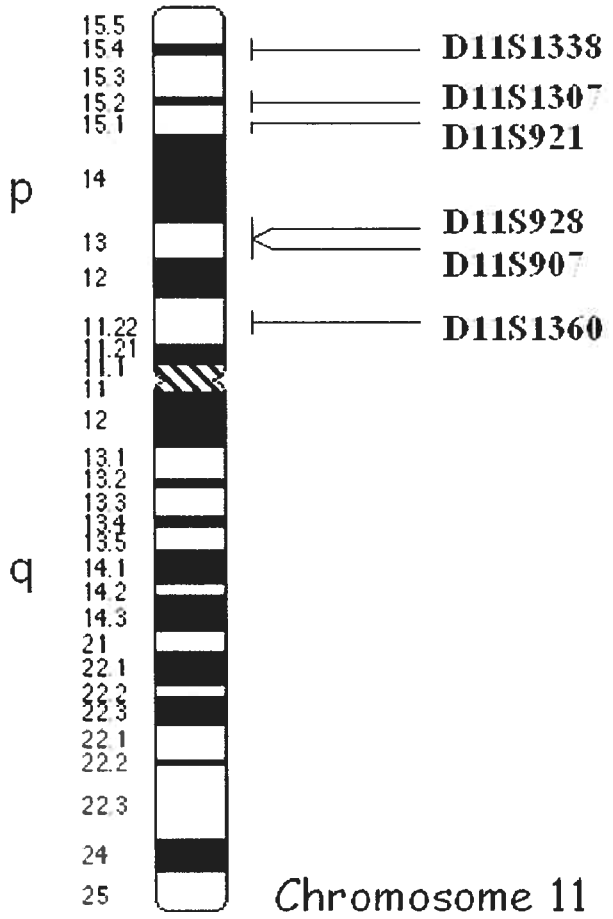
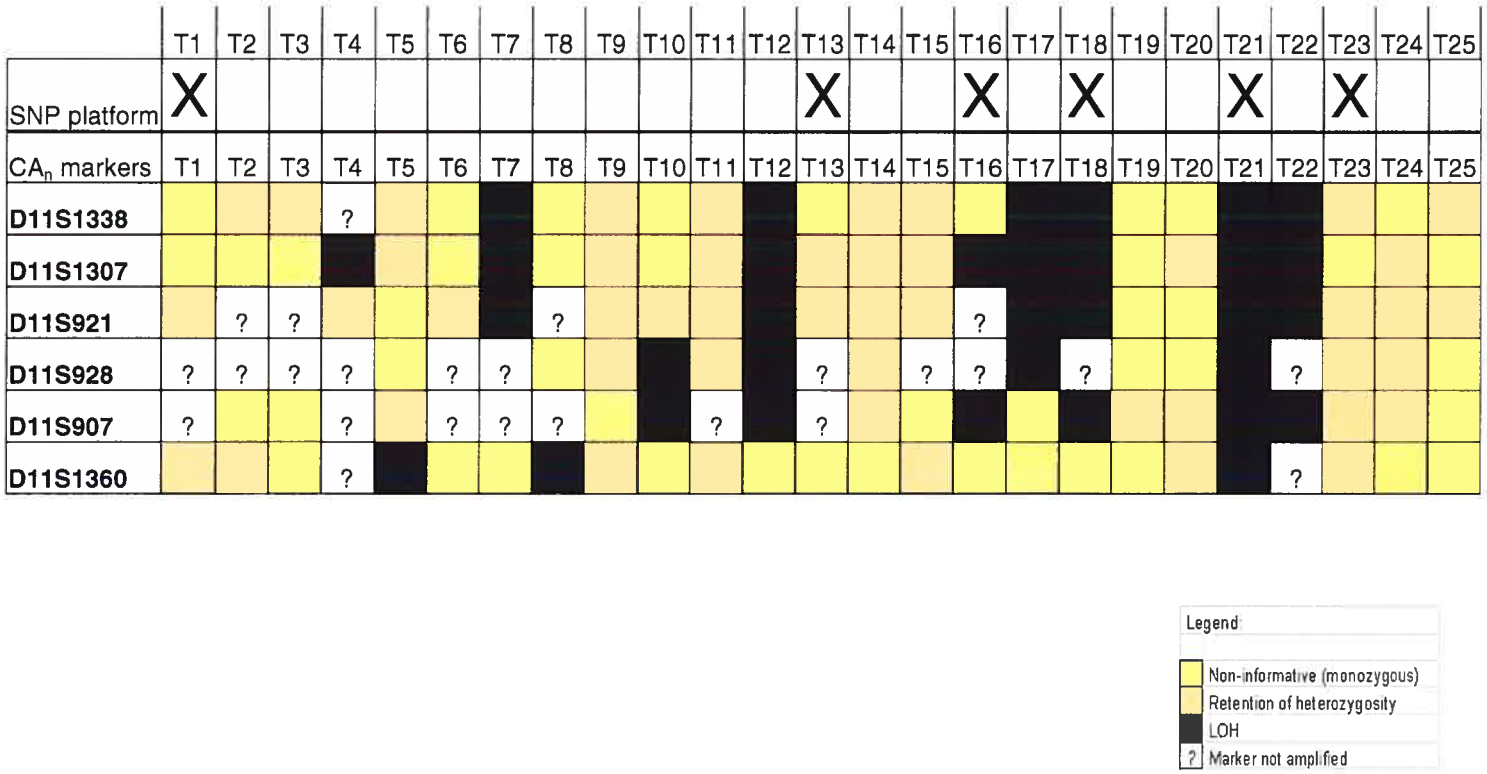


***Figure 9**: Microsatellite markers used in the LOH study with Wilms' samples. These markers were selected from the 1993-94 Genethon human genetic linkage map published in [66].*

b. Results

As expected [36], some Wilms' samples (11 out of the 25, representing 44%) showed a loss of heterozygosity in the regions 11p15 and 11p13. The detailed map of chromosome deletion for the 25 patients is illustrated in table VI.

***Table VI**: Chromosome deletion map after amplification of chromosome 11 markers in 25 Wilms' tumours. Some microsatellite markers did not amplify and are indicated by an interrogation mark (?). With the completion of the detailed map of the loss of heterozygosity, we selected 6 specific Wilms' samples (indicated by an X in the table) to test the Illumina platform.*



4. Tumoral genotyping with Illumina SNP array

In recent years a variety of studies have emerged measuring total chromosomal copy number at increasingly high resolution by using the Affymertix SNP platforms [36, 50, 54, 61, 67-69]. Since 2005, Illumina platform is available at Genome Quebec allowing easy access to data analysis for a lower price range.

Illumina Inc. currently manufactures several formats of high-density SNP genotyping arrays; the Sentrix Human-1 SNP BeadChip (109k SNPs, 70% of which are located in exons or within 10kb of transcripts), the HumanHap300 (317k tag SNPs), and a higher-density HumanHap550 (550k tag SNPs) [70]. The Illumina platforms were initially tested by using cancer cell line HL-60 human promyelocytic leukemia containing several well characterised chromosomal deletions and amplifications [71]. To our knowledge, these platforms were not yet used to examine aberrations in cancer samples obtained from patients. For this reason, it was important to test this new platform with our collection of Wilms' tumours before using it to study rare paediatric cancer.

4.1. Sentrix Human-1 SNP BeadChip array

The Sentrix Human-1 SNP BeadChip with a total of 109 000 SNPs was used for this study (illustrated in figure 10). The average spacing between SNPs on the Human-1 SNP BeadChip is 26 kb; median spacing is 13 kb. The Infinium WGG assay was employed for the high resolution analysis and profiling of both LOH and DNA copy number changes in constitutional and samples. The Infinium Whole Genome Genotyping (WGG) procedure is composed of four automated steps: (1) whole genome amplification, (2) hybridization to a specific and sensitive oligonucleotide probe array, (3) an array-based SNP scoring assay, and (4) signal amplification [72]. The procedure is illustrated in figure 11 and summarized in figure 13. The infinium protocol process minimizes the amount of time, sample volume and material required to carry out tumoral genotyping.

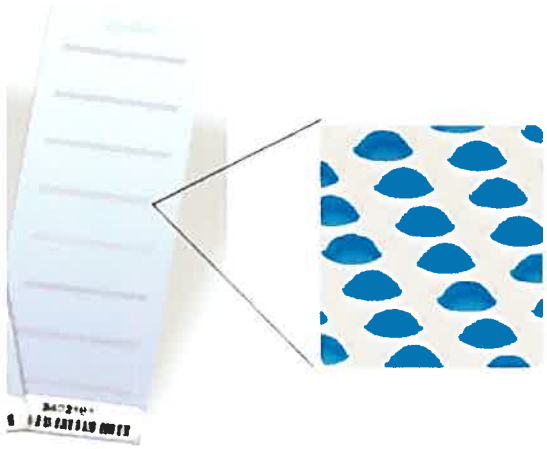


Figure 10: *The sentrix Human-1 SNP BeadChip allows processing of 8 samples at a time; cost of approximately 600$ per sample.*

4.2. New array-based whole-genome genotyping (WWG) assay

The DNA sample used for this assay is isothermally amplified overnight. Additionally, a relatively low DNA sample requirement (750 ng) is sufficient to assay 100,000 SNP loci. The amplified product is then fragmented by a controlled enzymatic process which does not require tedious gel analysis. After alcohol precipitation and resuspension, the BeadChip is prepared for hybridization in the capillary flow-through chamber; samples are applied to BeadChips and incubated overnight. The amplified and fragmented DNA samples anneal to locus-specific 50-mers (covalently linked to one of over 200,000 bead-types) during the hybridization step. One of two bead-types corresponds to each allele per SNP locus. After hybridization, allelic specificity is conferred by enzymatic extension known as allele-specific primer extension and the SNP scores are then analysed with BeadStudio visualization tool Illumina genome viewer (IGV).

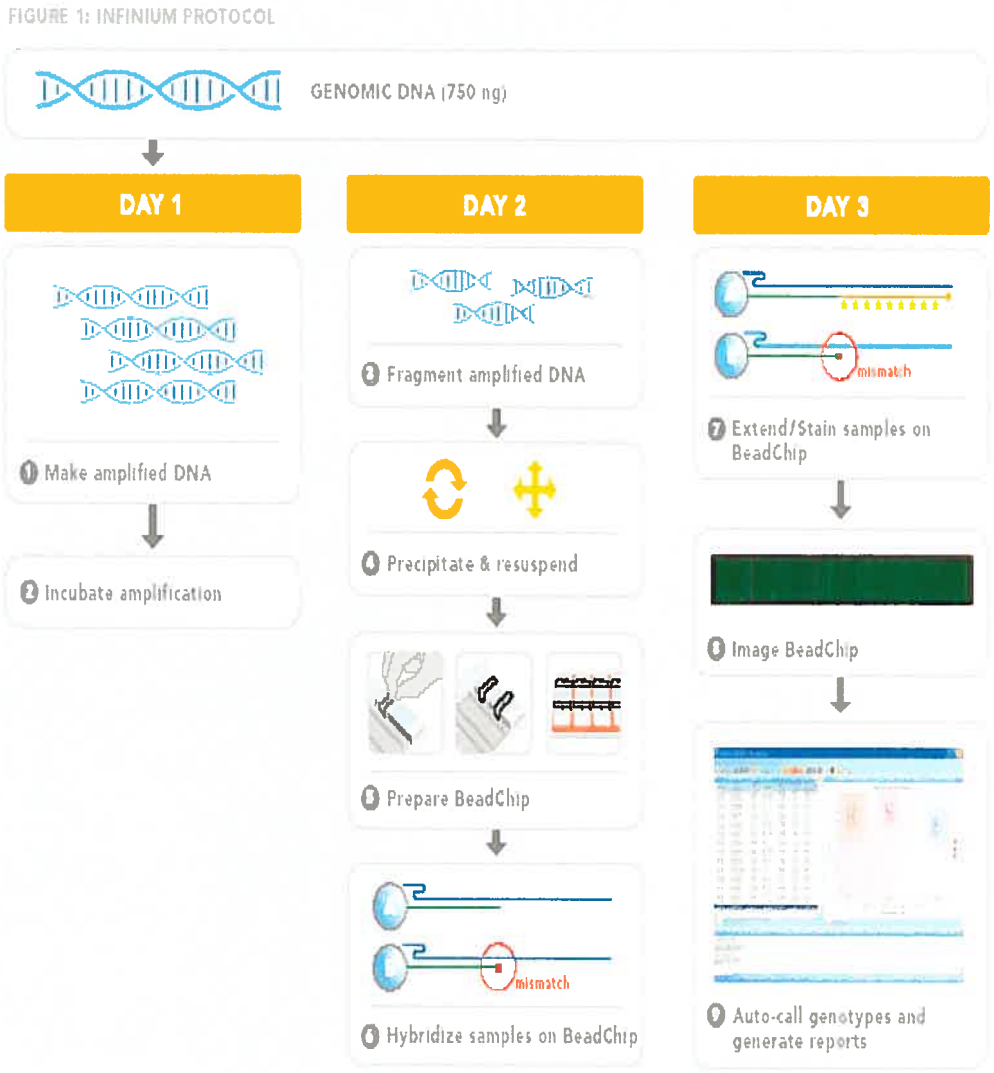


***Figure 11**: Illumina Work flow.*

***Steps 1 and 2**: DNA amplification.*

***Step 3:** Fragmentation of amplified product.*

***Step 4:** Alcohol precipitation and resuspension.*

***Step 5:** Preparation of the BeadChip.*

***Step 6:** Hybridization step.*

***Step 7:** Products are fluorescently stained*

***Steps 8 and 9:** Analysis of the auto call genotypes and generated reports.*

http://www.illumina.com/General/Products/SNP/pdf/INFINWKFLOW.pdf

4.3. Allele-specific primer extension (ASPE)

For the Infinium I assay adapted for the Sentrix Human-1 SNP BeadChip, two beads (A and B) for each SNP are used to score the SNP site. The probe sequence of these beads differs only at the 3' terminal base (opposition SNP site), creating an allelic discrimination in the polymerase extension. In this assay, known as allele-specific primer extension (ASPE), the perfect matched bead type will preferentially extend over the mismatched bead type and gets labelled (Figure 12) [72]. The genotype state of a given SNP locus (AA, AB, or BB) is determined by the intensity ratio between the two corresponding bead types. To do so, the BeadChips are scanned with a two colour confocal Illumina BeadArray Reader. The fluorophors generated during signal amplification/staining extension products are excited, the image intensities extracted, and the resultant data are analysed to determine SNP genotypes using Illumina's BeadStudio software [72].

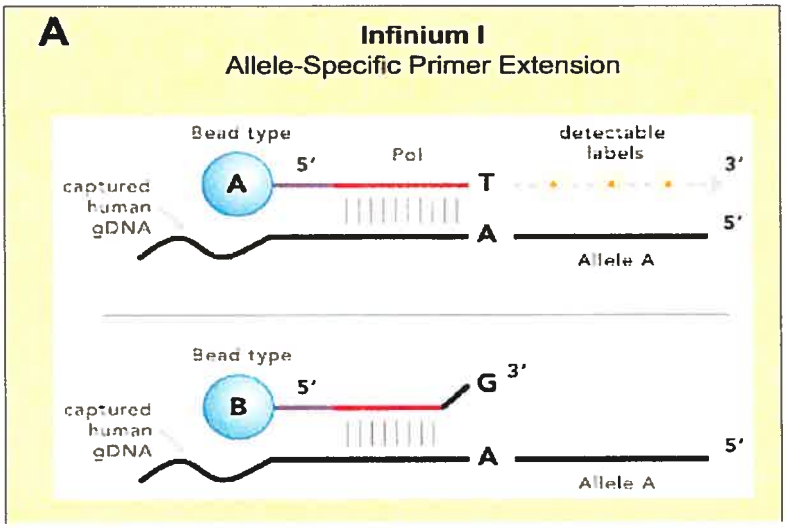


***Figure 12**: Infinium I is an ASPE-based one colour assay*

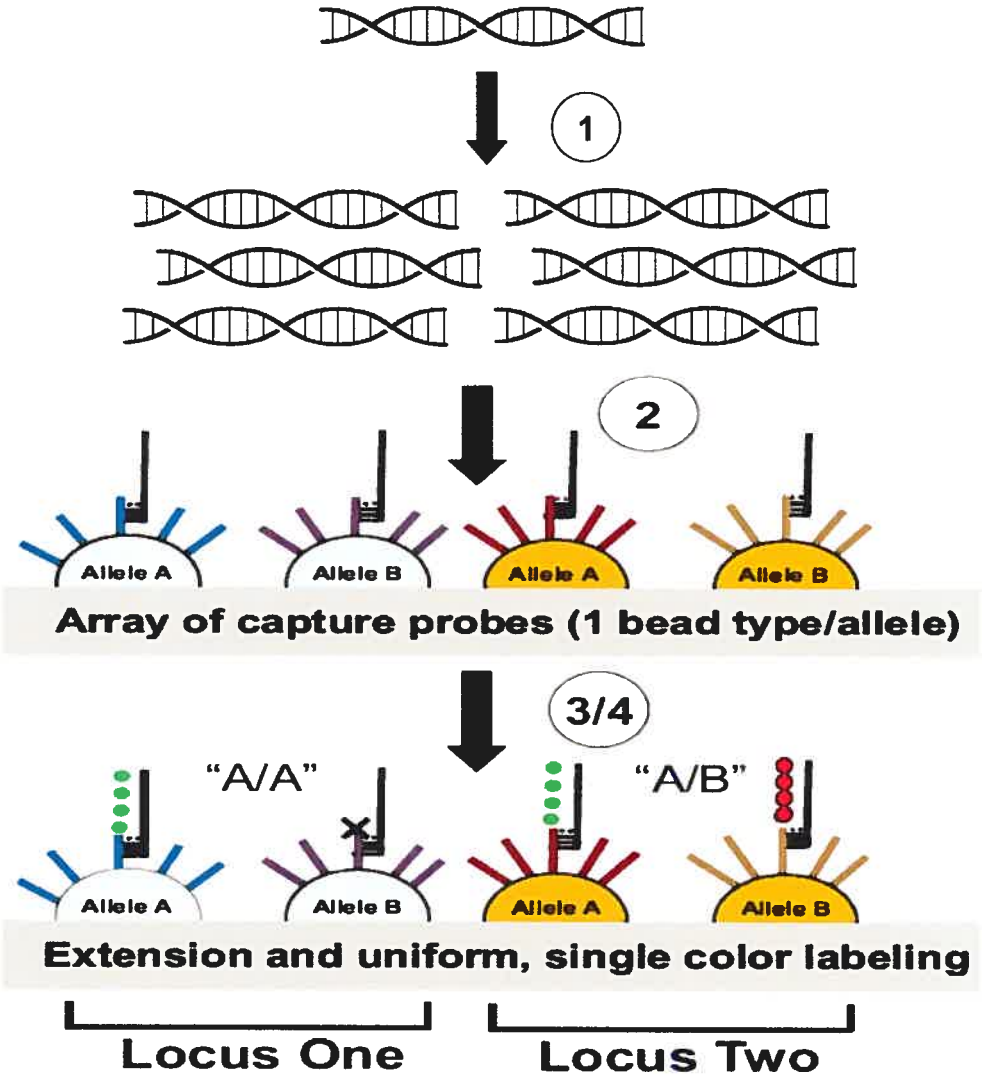


Figure 13: *Summary of the Infinium assay.*

The Infinium assay uses the complete complexity of the genome. After whole genome amplification, ASPE is used to score the SNP site, requiring two beads for each SNP.

4.4. Data visualization

To visualize the results, the data for a particular SNP is plotted as polar coordinates. The x-axis represents theta which is the angle deviation from pure A signal, where 0 represents pure A signal and 1.0 represents pure B signal, and the y-axis represents the distance of the point to the origin. Data are normalised by the GenTrain/GenCall normalization algorithm (figure 14). The latter removes outliers, adjusts for channel-dependent back-ground and global intensity differences, and also scales data [71].

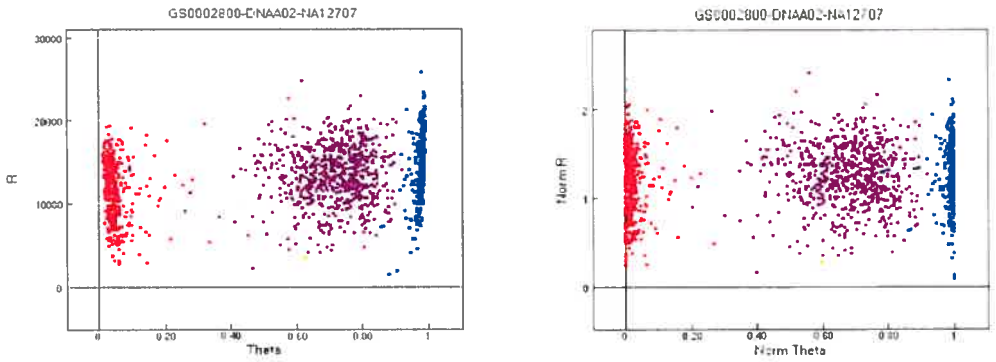


***Figure 14**: Normalization Turned Off & Normalization Turned On (© Illumina Inc. BeadStudio Genotyping Module User Guide Manual) The X and Y color channels undergo an affine coordinate transformation to make the data appear as canonical as possible with the homozygotes.*

4.5. BeadStudio visualization tool: Illumina genome viewer (IGV)

BeadStudio provides an easy way to analyze samples for Loss of Heterozygosity (LOH) and other chromosomal aberrations. The primary tool for LOH analysis is the Illumina Genome Viewer (IGV) permitting to view:

- B Allele Frequency (corresponding to the theta value for the SNP corrected for cluster position) permitting to view the genotype of the analysed sample,
- Log R Ratio to process the B allele frequency,
- LOH Score estimating the likelihood of a region exhibiting loss of heterozygosity,
- Copy Number Score estimating the copy number at an individual locus,
- Copy Number P-Value: provides values representing confidence levels in regions with a copy number change.

The first three tools are illustrated in figure 15. The algorithm examines the allele frequency data for aberrations and the log R ratio for changes in copy number (increase, decrease, or no change). Once the algorithm identifies a significant deviation in the allele frequency, the log R ratio is examined.

• A small decrease in the log R ratio is bookmarked as a **heterozygous deletion**.

• A large decrease in the log R ratio is bookmarked as a **homozygous deletion**.

• An increase in the log R ratio is bookmarked as an **amplification**.

• If there is no change in the log R ratio but a loss of heterozygotes in the allele frequency, the region is bookmarked as **copy-neutral LOH**.

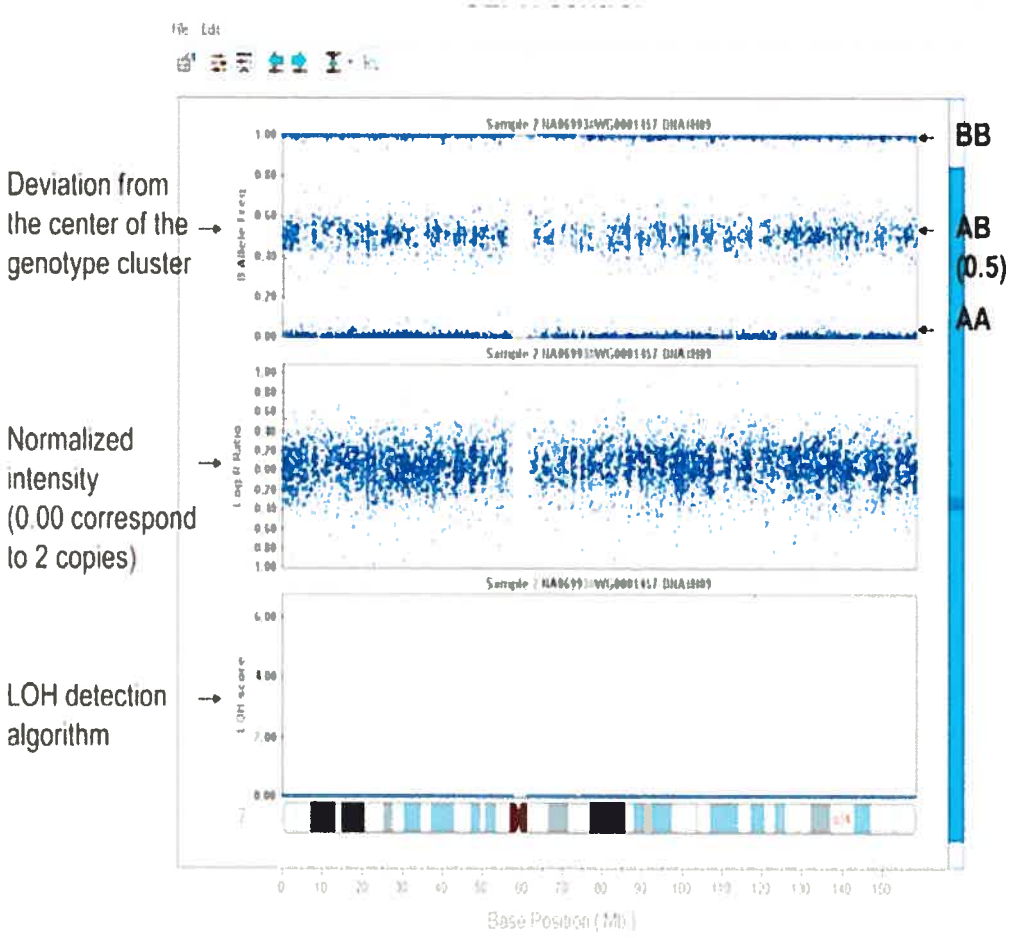


Figure 15: Illumina Genome Viewer (IGV). *The top window illustrates the B allele frequency where 0.0 represents AA genotype, 0.5 represents AB genotype and 1.0 represents BB genotype. The middle window represents the Log R ration where 0.0 corresponds to 2 copies. The bottom window illustrates the LOH detection algorithm where 0.0 corresponds to no LOH.*

Detection of chromosomal aberrations is accomplished by comparing the normalized intensity of a subject sample to a reference cluster sample(s). The reference values are derived from canonical clusters created from clustering ~120 normal reference samples obtained from the Centre d'Étude du Polymorphisme Humaine (CEPH) [71]. On the other hand, BeadStudio 2.0 loss of heterozygosity plus (LOH Plus) module permits a paired sample mode allowing direct intensity comparisons between a subject sample and its corresponding matched pair, for example normal and tumoral samples from the same patient. The LOH score is a measure of the likelihood that a SNP is exhibiting LOH around a window region (over all N SNPs where N is the number of SNPs in a user-designed window size) centered at the SNP's chromosomal position. The window size depends upon the density of SNPs on the platform used to analyse the samples.

5. Tumoral genotyping results

In total, 6 Wilms' samples were used to evaluate the ability of the 109K SNP array to detect single-copy losses on chromosome 11 and to validate the platform. By using the BeadStudio visualization tool Illumina genome viewer (IGV), we were able to detect both chromosome 11 retention of heterozygosity and chromosome 11 LOH which were first obtained by microsatellite marker amplification. Additional chromosomal aberrations not detected with the microsatellite amplification were also observed. This section illustrates interesting findings for all Wilms' samples analysed by SNP array:

a. LOH
    a. Chromosome deletion
    b. Regional LOH
    c. LOH result comparison between SNP array and microsatellite
b. Copy-neutral LOH
c. Chromosome Duplication
d. Regional Amplification
e. Possibility to assess LOH without paired samples
f. Deletion mapping of the samples analysed in this study

a. LOH

a. Chromosome deletion

The example illustrated in figure 16 demonstrates LOH on chromosome 16. The chromosome 16q deletion and partial 16p deletion (red bar) are detected by a deflection in the log R ratio and the collapse of heterozygotes in the allele frequency.

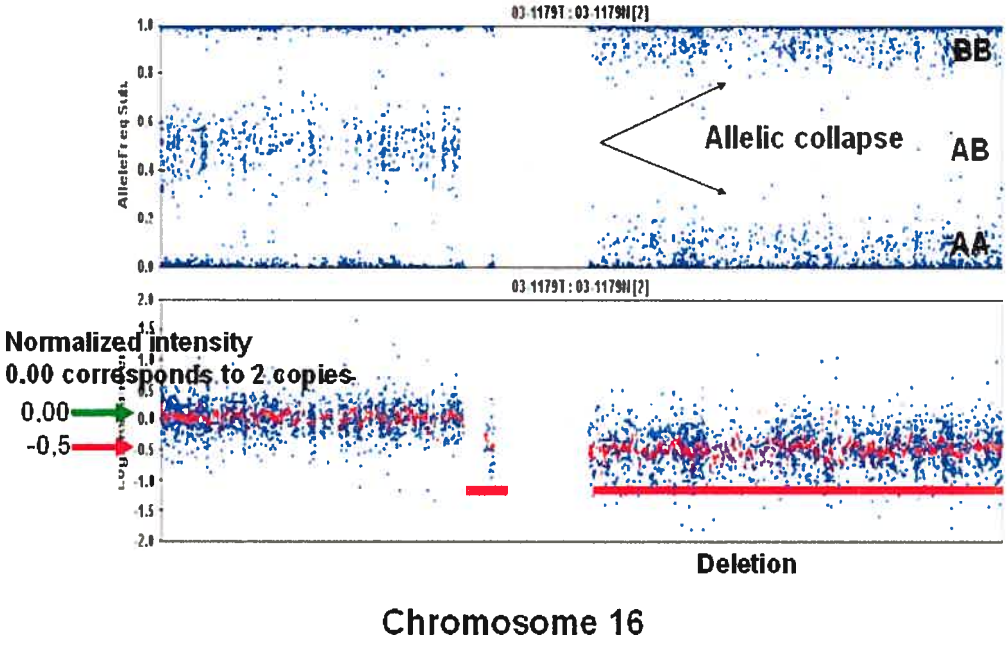


Figure 16: LOH. *Chromosomal deletion causing reduction to hemizygosity.*

The region localised between 35 Mb and 90 Mb on chromosome 16 in the tumour sample appears to show a loss of chromosome which is identified with high LOH scores, illustrated in figure 17.

Chromosome arm 16q is a common site of loss of heterozygosity (LOH) in Wilms tumors (WTs) [73, 74]. The 16q22.1 band harbors insulator protein CTCF gene, raising the possibility that reduced CTCF could lead to LOI of IGF2 in some cases. CTCF protein binds to DNA upstream of the H19 gene on chromosome band 11p15, and maintains normal imprinting of H19 and IGF2. Thus, its loss might predispose to de novo methylation of the maternal allele of H19 and loss of imprinting (LOI) of IGF2 in WTs [73, 74]. Other interesting associated cancer genes found on chromosome 16 include but are not limited to **FUS** (http://www.gene.ucl.ac.uk/nomenclature/data/get_data.php?hgnc_id=4010fusion (involved in t(12;16) in malignant liposarcoma) on the 16p11.2 sub-band, **RBL2** (retinoblastoma like-2) at 16q12.2, **MMP2** (matrix metallopeptidase 2) at 16q13-q21, **NQO1** (NAD(P)H dehydrogenase quinine 1 at 16q22.1, **ATBF1** (AT-binding transcription factor 1) at 16q22.3-q23.1, **FANCA** (Fanconi anaemia complementation group A) at 16q24.3. These informations can be found on the Atlas of genetics and cytogenetics in Oncology and Haematology (http://atlasgeneticsoncology.org).

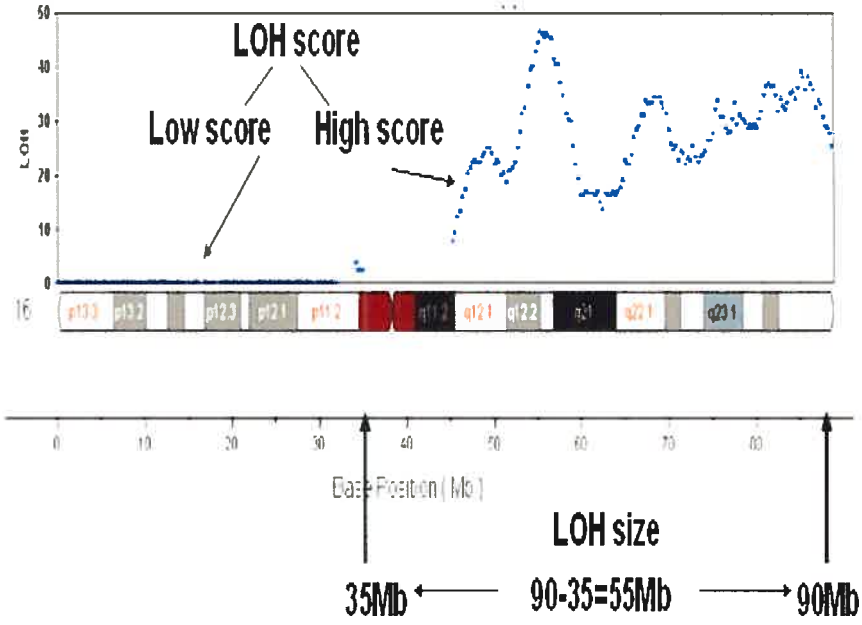


***Figure 17**: LOH size on chromosome 16p.*

This is another example demonstrating LOH, this time on chromosome 3 illustrated in figure 18. The chromosome deletion is again detected by a deflection in the log R ratio and the collapse of heterozygotes in the allele frequency. An interesting gene located at the sub-band 3p21.3 is **RASSF1** (Ras association (RalGDS/AF-6) domain family 1) implicated in many caners including clear cell renal carcinoma [75], lung cancer [76]; ovarian and renal cell carcinoma [77]; bladder carcinoma [78], nasopharyngeal carcinoma [79].

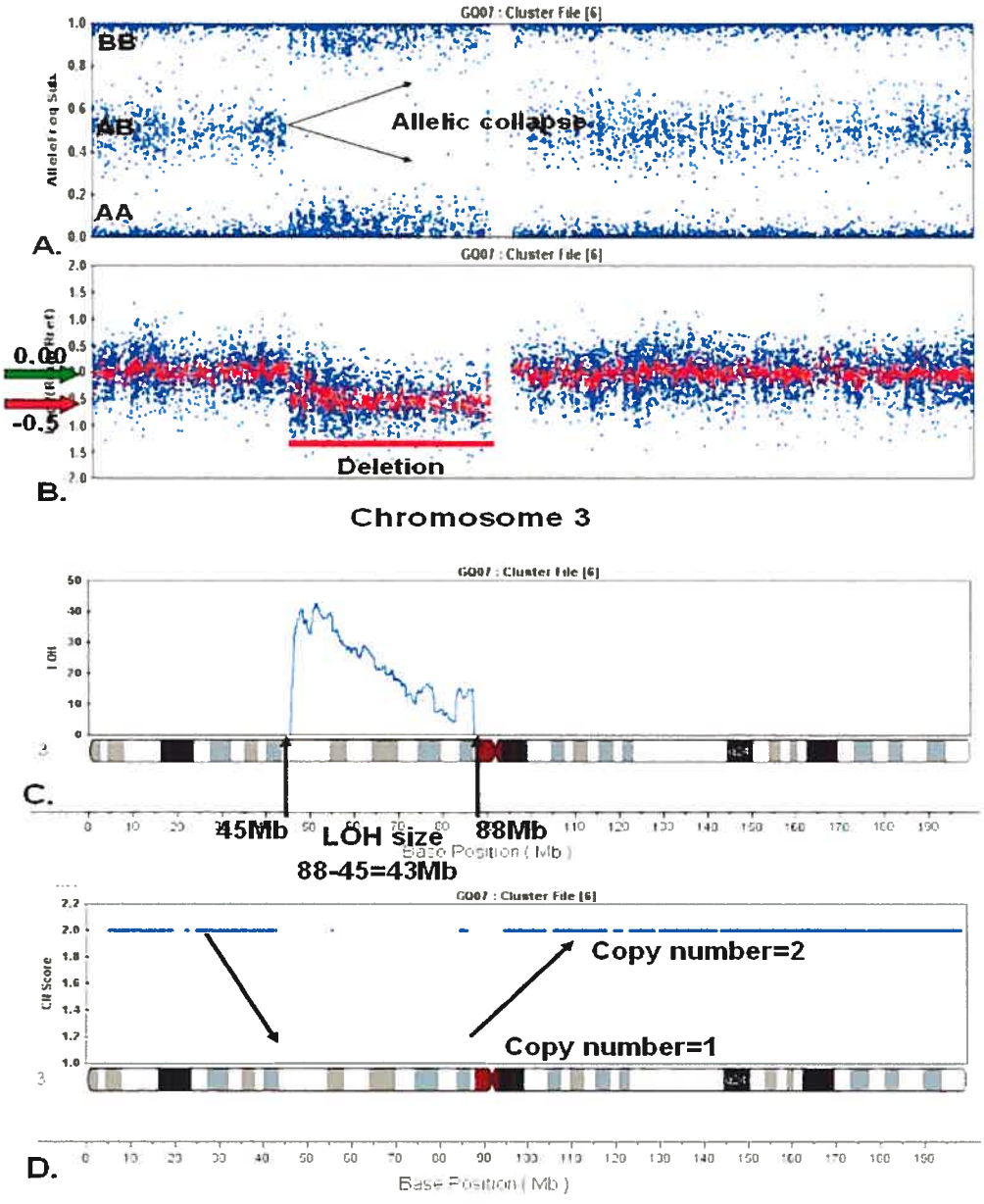


***Figure 18**: Chromosome 3 deletion represented by Allele frequency (window A), Log2 deviation (window B), LOH high score (window C) and Copy number change (window D).*

The copy number Bonferroni-adjusted P-value algorithm returns a value between 0 and 1. A value of 1 signifies no change in the copy number state of a given locus. The closer the value gets to 0, the higher the likelihood that there is a change in copy number for the sample. In this example, the whole chromosome 7 is deleted (figure 19) for which the copy number p-value gives a number of 0 throughout the whole chromosome, indicating a high probability of LOH (figure 20). Chromosome 7 deletions have been described in Wilms tumours including chromosome 7p14 germline mutations of the POU6F2 gene (the POU domain, class 6, transcription factor 2; also known as RPF1) [80-84].

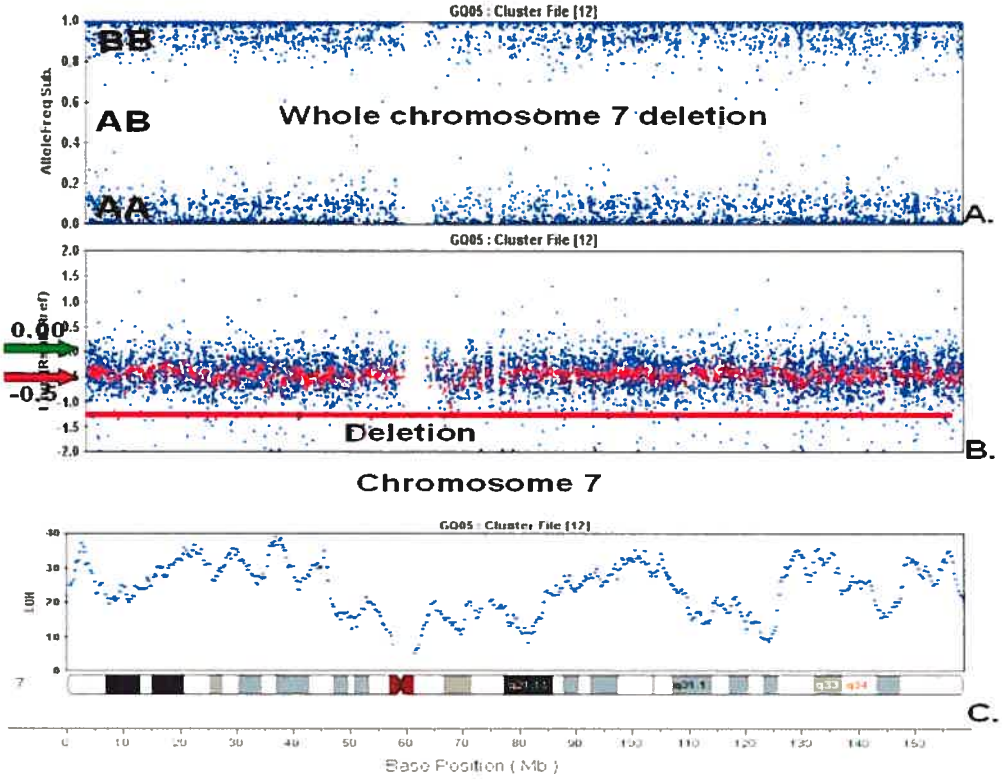


***Figure 19**: Whole chromosome 7 deletion represented by Allele frequency (window A), Log2 deviation (window B), LOH high score (window C)*

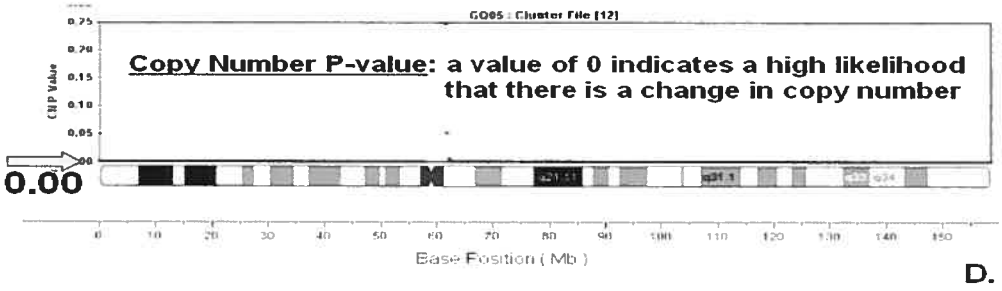


***Figure 20**: Copy number p-value of 0 (illustrated in window D) throughout the whole chromosome 3 indicates a high probability of copy change number.*

b. Regional chromosome deletion

A SNP of interest with its location can be viewed with Illumina Chromosome Browser (ICB) permitting to associate a chromosomal aberration with a specific SNP panel. All the available genes for the selected window can be viewed by ICB. In addition, all gene exons, gene strand type, transcription start/end positions, and coding region start/end positions for a selected gene can be viewed in the gene detail window. This tool gives access to a tremendous amount of information permitting to localize deleted or amplified genes in the samples analysed by Illumina Sentrix Human-1 SNP BeadChip.

In one case (sample 13 illustrated in figure 21), Illumina Chromosome Browser shows an interesting chromosome region deletion at 9q22.32 containing many interesting genes including PTCH gene. Somatic mutations in the PTCH2 gene (OMIM: 603673) have been identified in basal cell carcinoma (OMIM: 605462) and in medulloblastoma (OMIM: 155255), both of which are features of the nevoid basal cell carcinoma syndrome. This gene was never before associated to Wilms'.

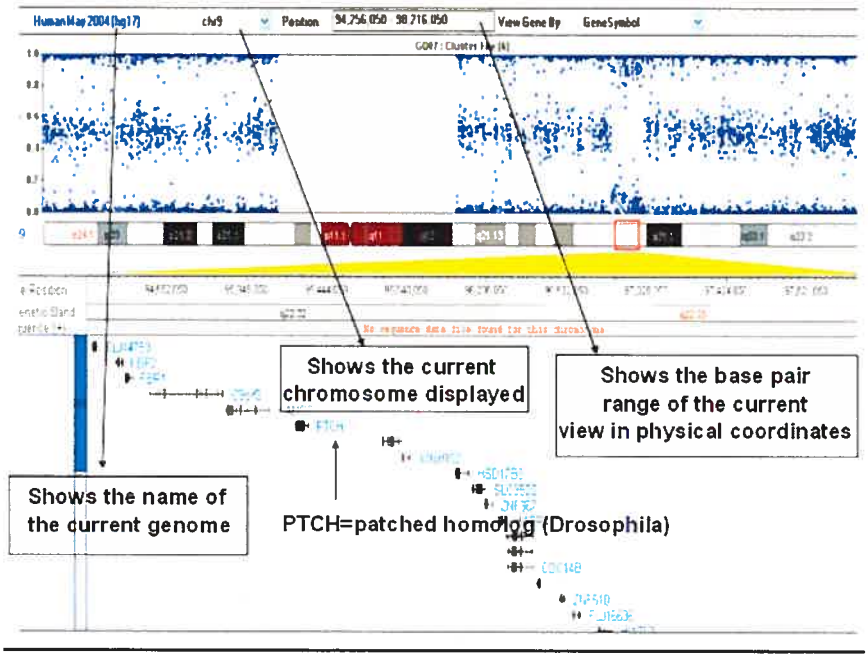


***Figure 21**: Illumina Chromosome Browser shows an interesting chromosome region deletion on 9q22 containing the PTCH gene.*

c. Comparison of the results obtained for the LOH study by microsatellite markers amplification versus SNP array

Finally, to test the concordance of the Sentrix Human-1 SNP BeadChip with the PCR-based microsatellite analysis, results obtained from CA_n amplification and SNP array were compared. Upon analysis, all information on chromosome 11 obtained by Sentrix Human-1 SNP BeadChip correlated with results obtained with microsatellite amplification except for tumour number 1. The results are shown in table VII.

The first microsatellite analysis of tumour number 1 did not demonstrate LOH on chromosome 11. For this reason, we wanted to re-do the microsatellite analysis a second time to verify the Illumina result. The sample was amplified with chromosome 11 markers D11S921 and D11S1360 and compared to its matching normal sample a second time. This time, the results revealed LOH on chromosome 11 as per Sentrix Human-1 SNP BeadChip demonstrated upon analysis. The PCR product with markers D11S921 and D11S1360 was misread the first time.

This proves that when the two methods were compared for chromosome 11, the SNP array is very robust and gives concordant results with the LOH study.

***Table VII**: LOH results comparison between SNP array and microsatellite amplification*

	CA						SNP
	1338	1307	921	928	907	1360	Short arm of Chr 11
T1	?	?	NO LOH	?	?	NO LOH	LOH
T13	?	NO LOH	NO LOH	?	?	NO LOH	NO LOH
T18	LOH	LOH	LOH	?	LOH	?	LOH
T21	LOH	LOH	LOH	LOH	LOH	LOH	LOH
T23	NO LOH	?	NO LOH	NO LOH	NO LOH	NO LOH	NO LOH

b. Copy-Neutral LOH:

LOH on chromosome 11p is illustrated in figure 22 is noted by loss a of heterozygosity in the allele frequency window. This is a copy-neutral LOH since there is no change in the log R ratio. Copy-neutral LOH does not produce a net reduction in chromosomal copy number; rather, it leads to loss of one set of parental alleles and duplication of the other.

Chromosome 11p contains the *IGF2* gene an imprinting-sensitive gene, which is active only on the paternal homologue. Due to the loss of **maternal** and reduplication of **paternal** chromosome 11p15.5, *IGF2* gene undergoes a 2-fold increase in functional gene dosage. In a previous study, loss of heterozygosity of 11p15.5 and loss of imprinting of IGF2 were reported as the most frequent genetic (29%) and epigenetic (40%) alterations in Wilms' tumours, respectively [85].

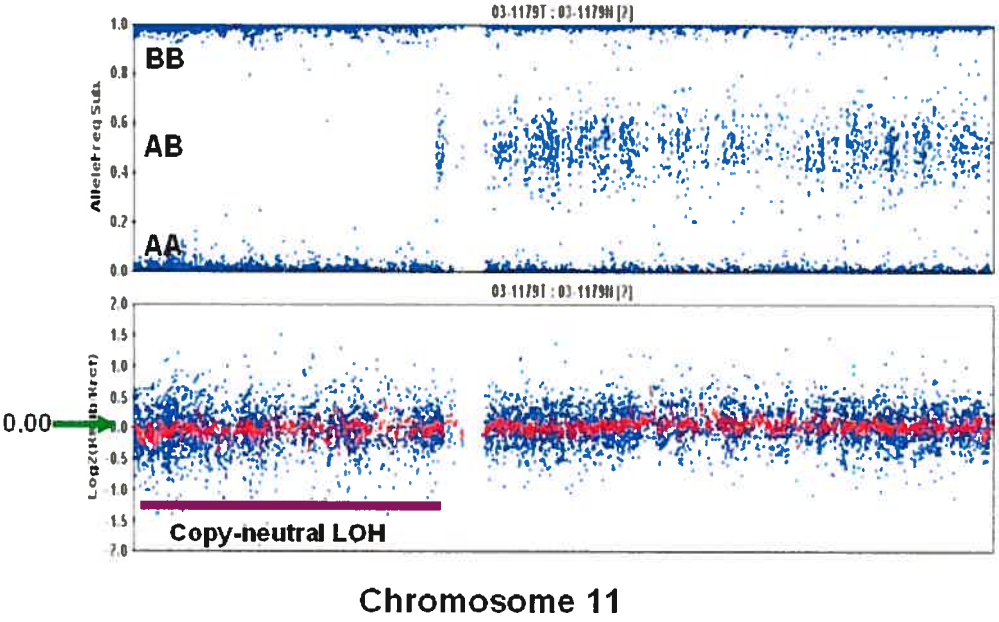


***Figure 22**: Copy-neutral LOH on chromosome 11p.*

c. Chromosome Duplication:

In figure 23, chromosome 19q is duplicated (black bar), inferring a total copy number of 3. Notice the increase of the log R ratio from 0.0 to 0.5 and the cluster split in allele frequency. In an amplified region, there may be four or more modes. The location of the modes indicates the amount of amplification. Here, the modes occurring near 0.33 and 0.66, a copy number of 3 is indicated (0 for AAA, 0.333 for AAB, 0.666 for ABB and 1 for BBB).

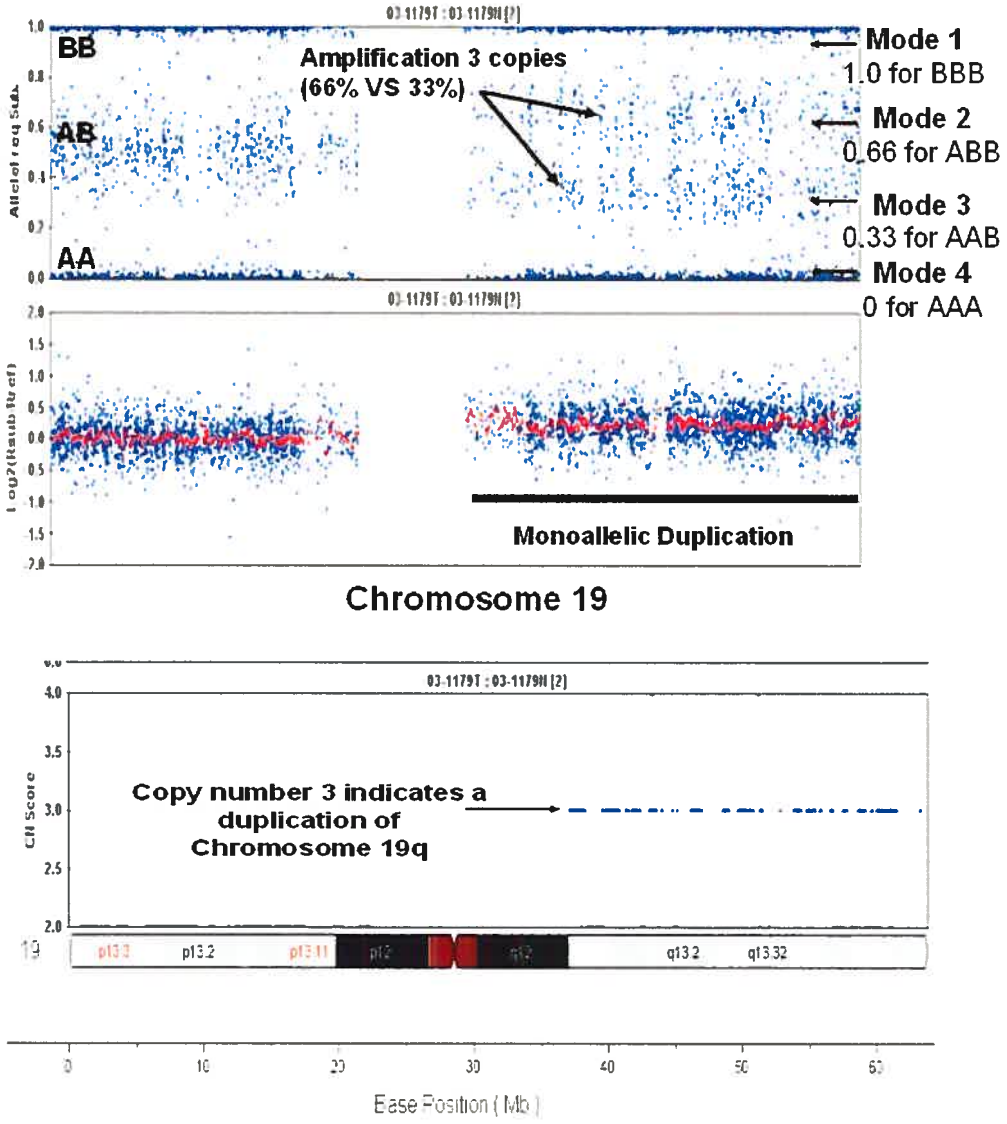


***Figure 23**: Chromosome 19q amplification. A copy number of 3 indicating a monoallelic duplication creating three copies of chromosome 19.*

In figure 24, the modes occurring near 0.25 and 0.75 show a copy number of 4 (0 for AAAA, 0.25 for AAAB, 0.75 for ABBB and 1 for BBBB). In this case, chromosome 4 is present in four copies.

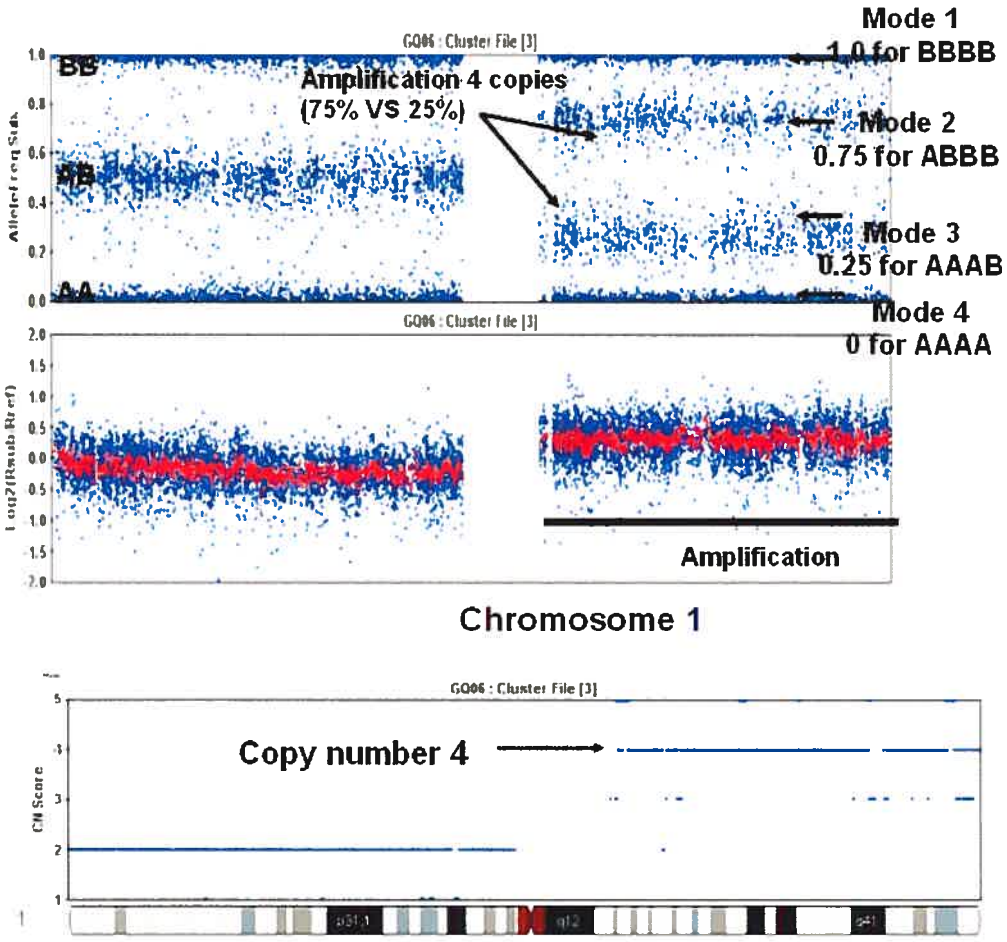


Figure 24: *Chromosome 1 amplification. Chromosome 1q is amplified with a copy number of 4 indicated by the 4 different modes: 0.0 for AAAA, 0.25 for AAAB, 0.75 for ABBB, 1.0 for BBBB.*

d. Regional amplification:

The monoallelic amplification of chromosome 1p36.11 to 1p35.2 is evidenced by an increase in the log R ratio and the split in the allele frequency (figure 25). The

zoom plot allows a detailed view of the interest chromosome region which is here illustrated in figure 26 a) and b). A better view of the allelic imbalance can be viewed in this case of the sample and not the normal sample when observing region 1p35-36.

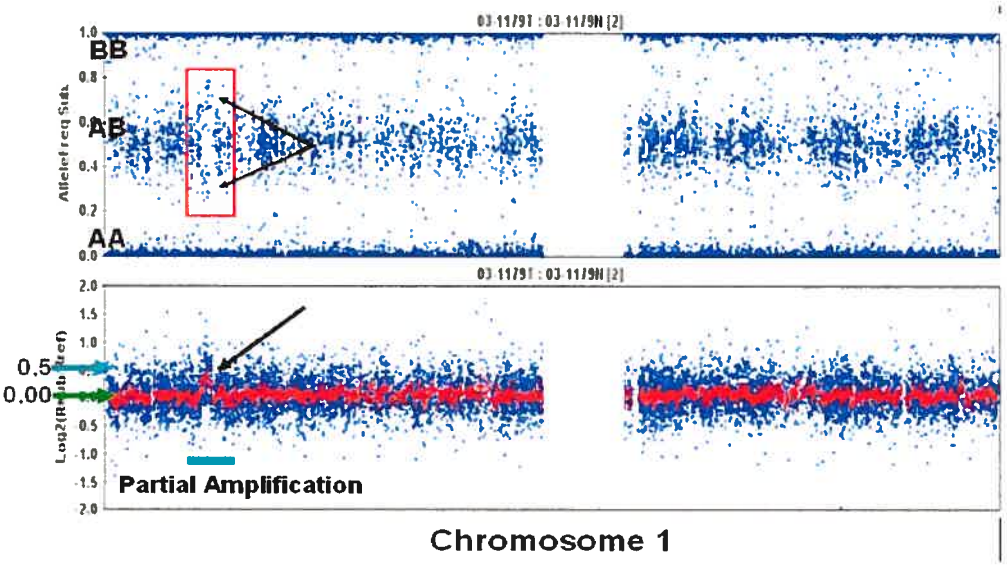


***Figure 25**: Amplification of chromosome 1p36.11 to 1p35.2*

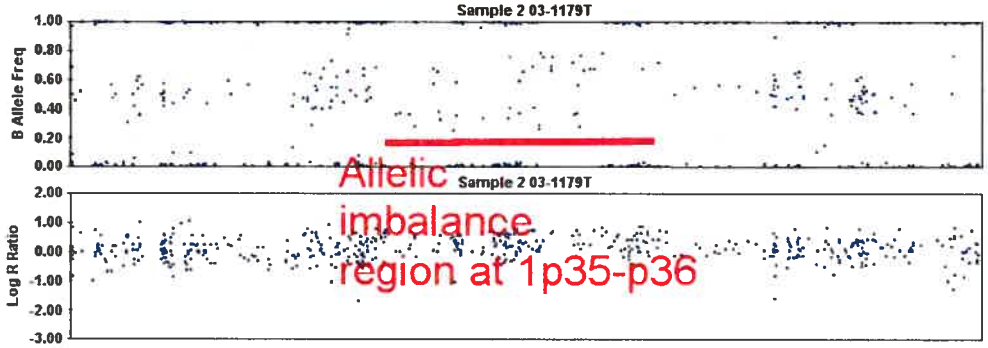


***Figure 26 a)**: Navigating the illumina chromosome browser. The zoom plot allows a detailed view of a particular chromosomal region. In this case, due to the small amplification of region 1p36.1-35.3 in the tumour sample, an allelic imbalance can be observed.*

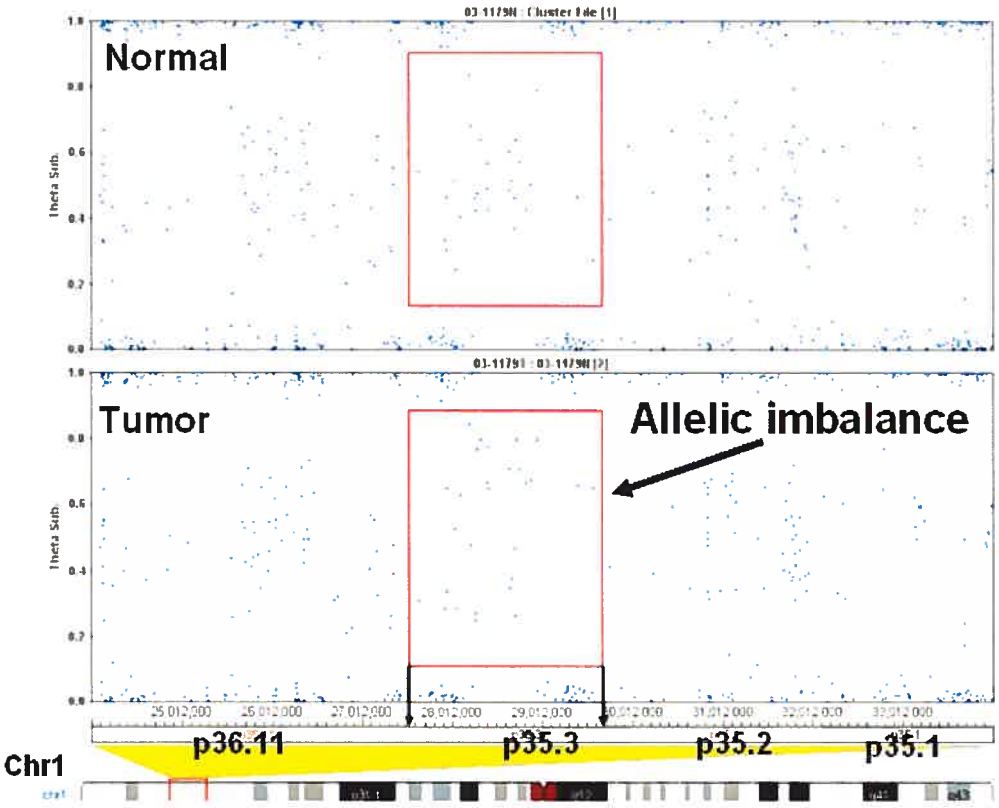


***Figure 26 b)**: Navigating the illumina chromosome browser, part 2. In this case, due to the small amplification of region 1p36.1-35.3 in the tumour sample, an allelic imbalance can be observed in the sample which is not observed in the normal sample.*

e- Possibility to assess LOH without paired samples

LOH studies by restriction fragment length polymorphism (RFLP) and by PCR-based microsatelite analysis require both normal and tumour tissue samples which are not always possible to obtain (patient has passed away or patient can not be retraced). The possibility to assess LOH without comparing a tumour sample to its matching normal sample would be a great achievement.

To assess the ability of the SNP platform to detect chromosome aberrations without the use of the matching normal sample, we compared the tumoral sample to its matching normal sample and the tumoral sample with the reference cluster by using LOH Plus Micorsoft Excel sample sheets (figure 27 a) and b)).

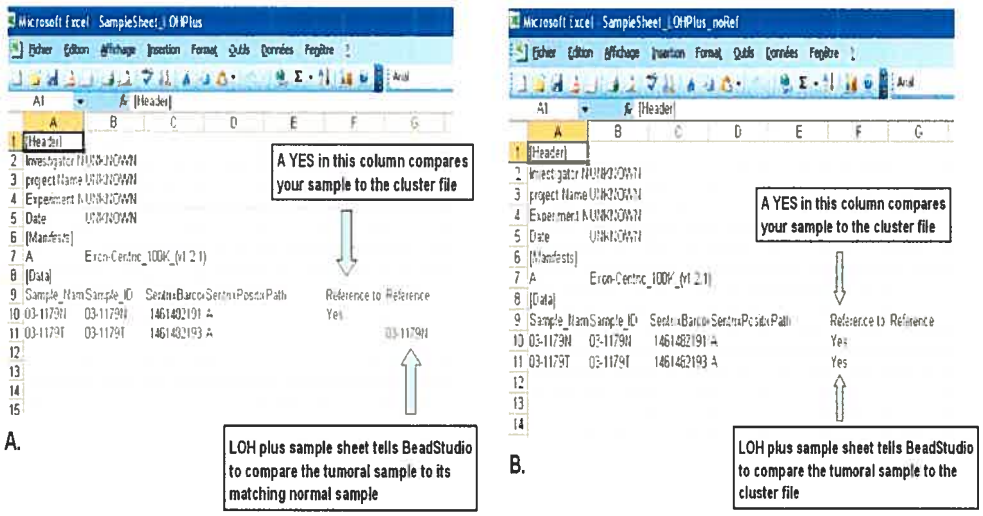


***Figure 27**: **A** Tumoral sample 21 was compared to its matching normal sample*
***B** Tumoral sample 21 was also compared to the reference cluster*

Using the 109K BeadChip, we successfully detected all of the aberrations previously detected on chromosomes 1, 11, 16 and 19 by comparing the tumoral

sample with its matching normal sample (figure 28) and with the Illumina cluster reference file (figure 29). This indicates that detection of chromosomal abnormalities can be achieved without necessarily comparing it to its matching normal sample. This is an important advantage when studying rare paediatric cancers.

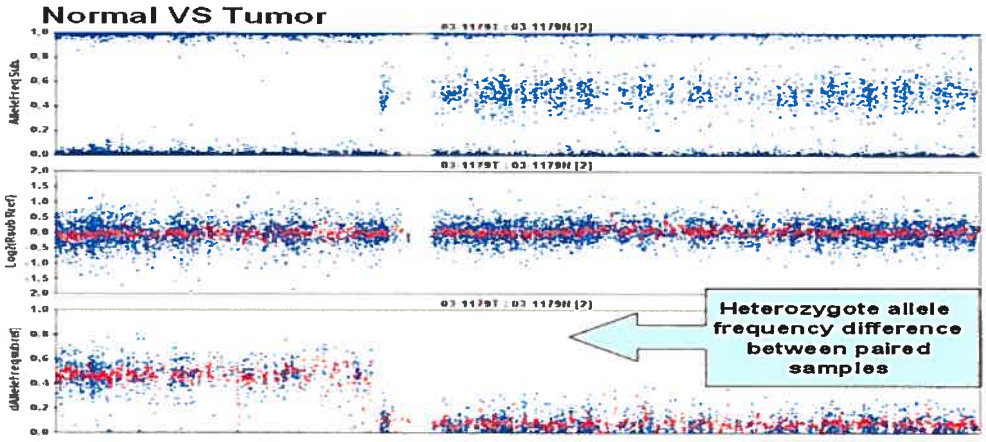


***Figure 28**: Tumour sample compared to its normal sample revealing chromosome 19q duplication.*

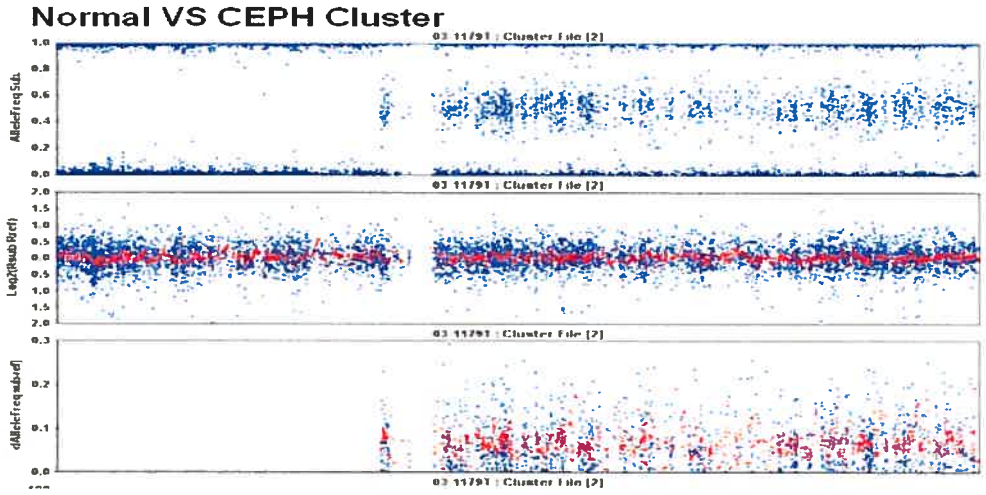


***Figure 29**: Tumour sample compared to the Illumina reference cluster revealing the same chromosome 19q duplication. However, the allele frequency ($Fre_{sub\text{-}ref}$) calculation illustrated in the third window shows more background noise when tumoral sample is compared to the reference cluster and does not give a value in the deletion region, even with a window scale change.*

e. Deletion/amplification mapping of the samples analysed in this study

A map of the smallest regions of deletion/amplification for the analysed Wilms' tumours was generated. The smallest region of deletion/amplification map allows the localization of oncogenes and tumour suppressor genes. However, manual mapping is time-consuming and is not as precise as a computerized mapping. Since each SNP can be precisely located in a deleted or amplified chromosomal region by Illumina Chromosome Browser, gene mapping would be greatly enhanced if the information obtained could be systematically incorporated into a database. These maps are generated using symbol bars illustrated in figure 30. The results we obtained from genotyping 6 Wilms' tumours are summarised in figure 31.

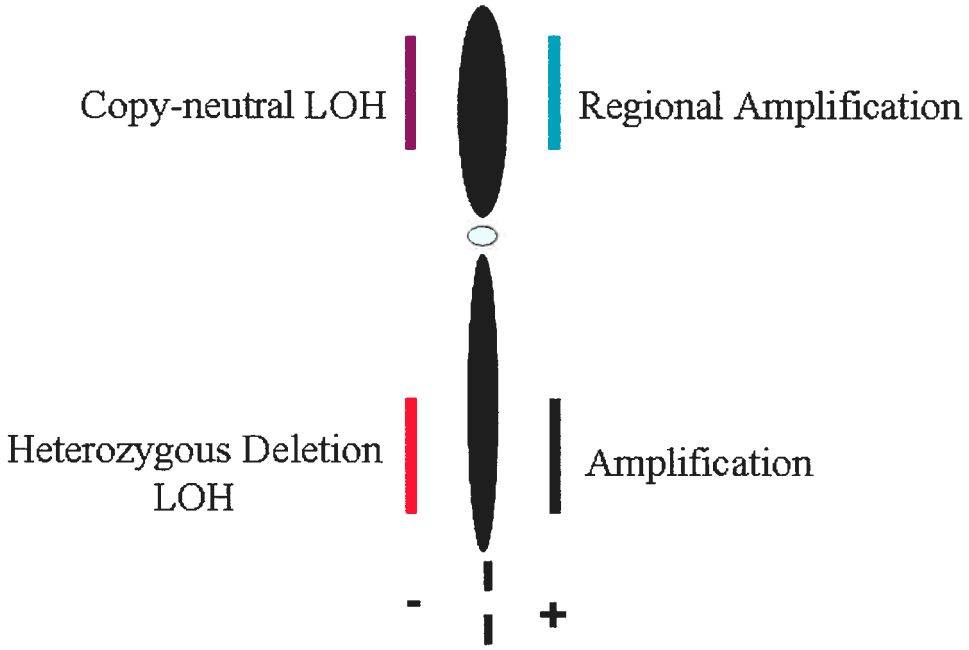


***Figure 30**: Legend for chromosome mapping: copy neutral deletions are represented by purple bars, heterozygous deletions are represented by red bars, partial amplifications are represented by green bars, amplification are represented by black bars.*

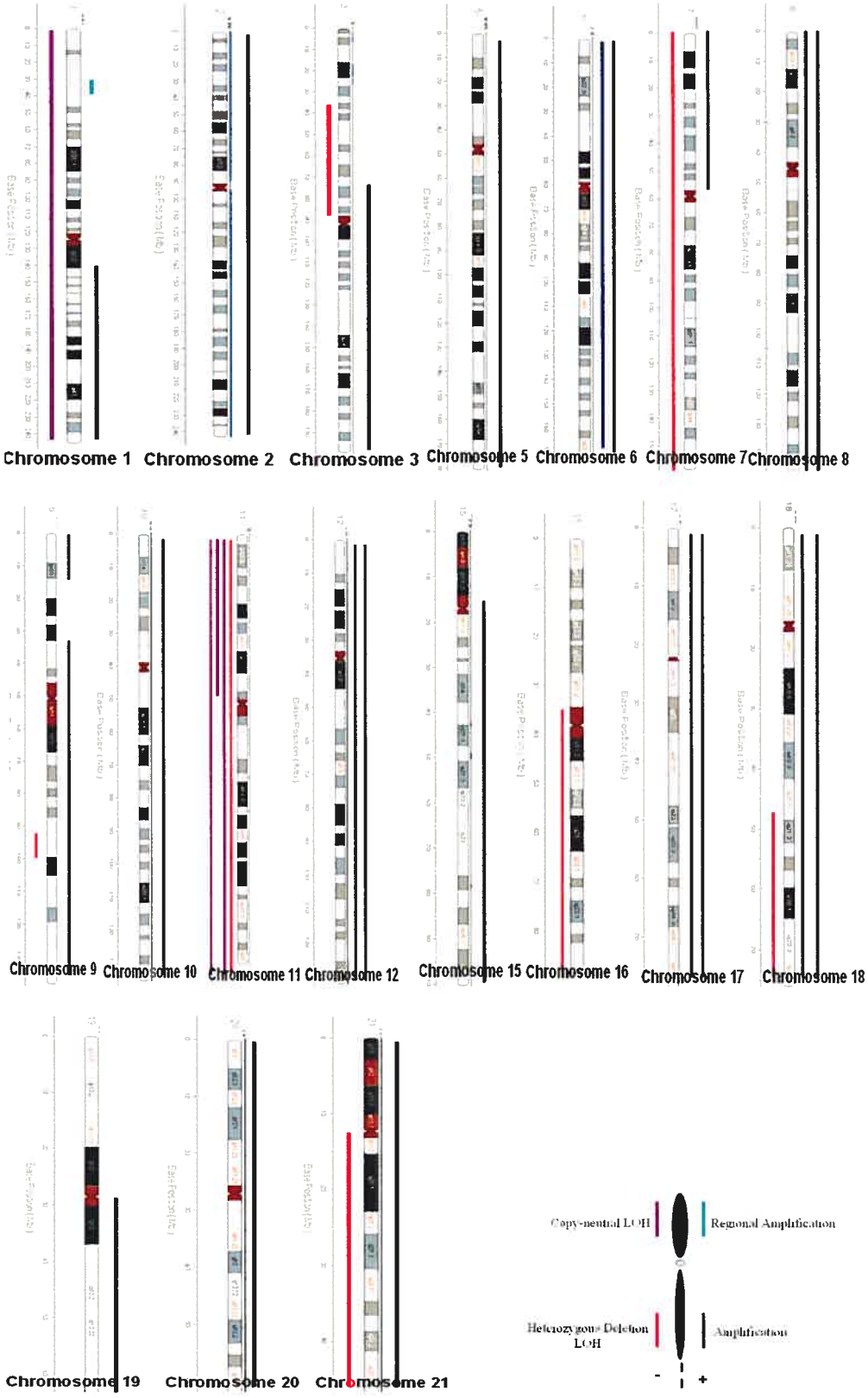


***Figure 31**: Chromosome mapping with Wilms' samples analysed on the SNP array.*

2 Evaluation of techniques in two orphan tumors

B. Rational for Infantile Myofibromatosis

We chose to study this tumour since there are a limited number of articles in the literature describing its molecular basis. Due to the limited information available on this tumour, we classified it as an orphan tumour. We applied various techniques to gain genetic information and to determine which methods are best suited to study orphan tumours:

- Familial study by linkage analysis,
- Somatic study was carried out by:
  - Microsatellite amplification of:
    - Frozen tumour samples
    - Microdissected paraffin embedded tissue,
  - Fluorescent in Situ Hybridization
  - Sentrix Human-1 SNP BeadChip array.

The results we have obtained are discussed in each section in order to illustrate the possible applications and limitations of each technique. These results also provide an overall idea of the future work that have to be achieved in the study of orphan tumors.

1.Introduction

Infantile Myofibromatosis (IMF) although an orphan paediatric cancer, is one of the most common fibrous soft tissue proliferations of infancy [86]. These lesions are present at birth and thus congenital and most are diagnosed during the first 2 years of life. Two types can be distinguished [87]. The solitary type, called infantile myofibroma, is defined by the presence of one nodule in the skin, muscle, bone or subcutaneous tissue. The multicentric type, called infantile myofibromatosis, can be further divided into two sub-types. In the first sub-type, the lesions are multicentric but without visceral involvement, while in the second, visceral involvement of the lung, gastrointestinal tract, liver and heart can occur [88]. Morbidity and mortality is associated directly with visceral involvement [89, 90]. Despite the fact that IM is the most common fibrous proliferation in infancy, many of its biological aspects remain unclear. In particular, IMF histogenesis is still not fully understood, although immunohistochemical staining and ultrastructural features suggest a myofibroblastic origin [91].

Review of the literature identified both sporadic and familial forms of IMF. The latter mostly follow an autosomal dominant inheritance [92-96] but there has been one published case of an autosomal recessive inheritance [97]. The etiology of IMF is still unknown. Cytogenetic investigations in two cases of solitary myofibromatosis showed a deletion del(6)(q12q15) in metaphase cells grown in vitro [98] and another case of monosomy 9q due to an unbalanced chromosome translocation t(9;16) (figure 32) [91] as the sole cytogenetic abnormalities. The familial cases occur earlier, are frequently neonatal, and multifocal [92-97]. These observations could indicate that IMF is caused by the loss of function of a yet unknown tumour suppressor gene. This is also supported by the autosomal dominant inheritance mode in some familial cases where the tumour is more aggressive. We choose to study this tumour as a model of orphan tumour since there is no molecular data information published in the literature on IMF.

Our objective was to quantify the percentage of Infantile Myofibomatosis with chromosomal deletion 6 and 9 in multiple samples obtained from CHU Ste-Justine and to find additional chromosome aberration in the sample genome.

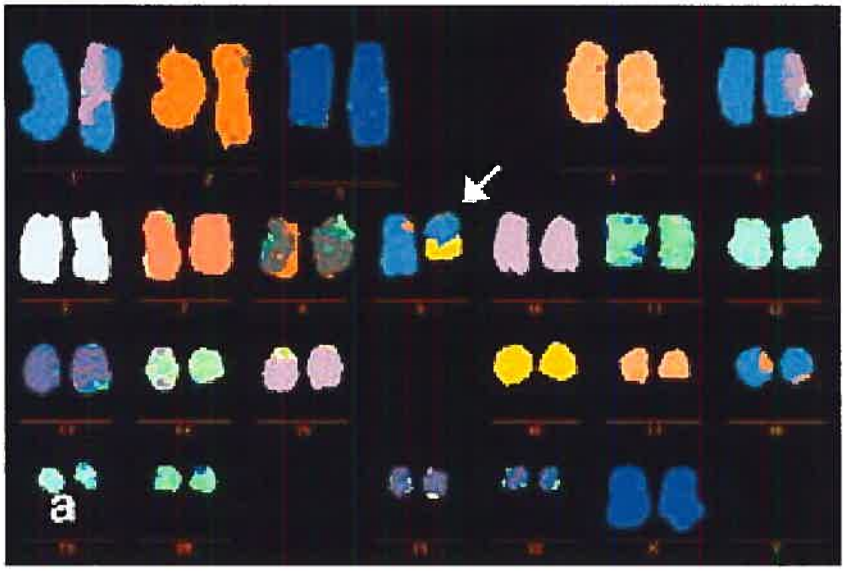


***Figure 32**: Multi-fluorescence in situ hybridization (FISH) analysis of a metaphase cell in a patient with Infantile Myofibromatosis showing monosomy 9q due to an unbalanced translocation t(9;16) as the sole chromosomal abnormality (arrow) [91].*

2. Linkage analysis

Since linkage analysis helped to clone many hereditary cancer genes over the years including tumour suppressor genes, we prerformed a linkage analysis with Infantile Myofibromatosis tumour samples. The linkage analysis was done in collaboration with Mount Sinai Hospital in New York. Blood samples were collected from the patients with a familial history of Infantile Myofibromatosis. DNA was isolated from the blood samples and sent to Mount Sinai hospital for analysis. We recruited two families followed at CHU Sainte-Justine with a history of Infantile Myofibromatosis (illustrated in figure 33 and 34). As of now, Mount Sinai hospital has a cohort of 10 families with a history of Infantile Myofibromatosis and it is still recruiting to increase the number of families for the familial study.

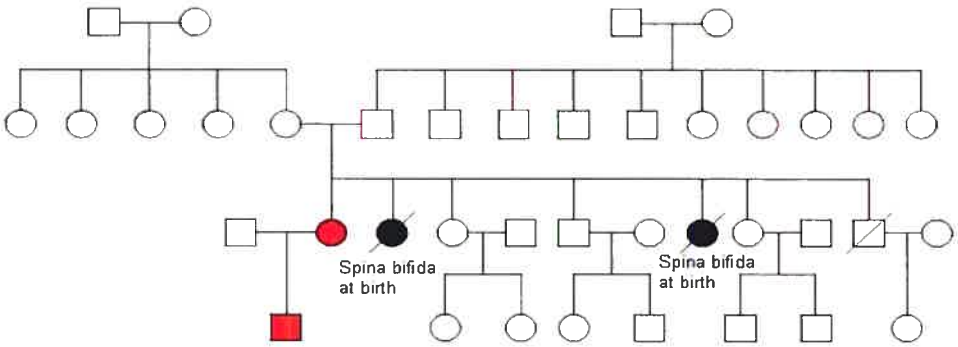


***Figure 33**: First family recruited at CHU Sainte-Justine with a history of Infantile Myofibromatosis (shown in red) used for the linkage analysis.*

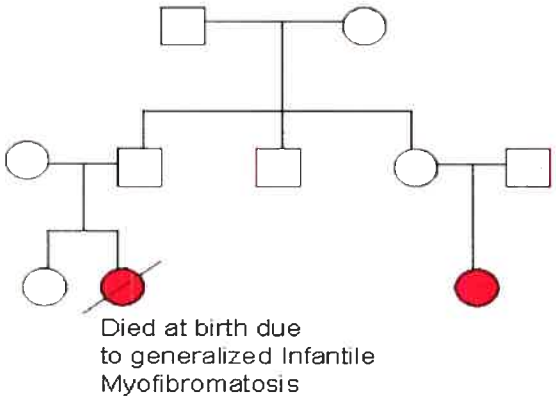


***Figure 34**: Second family one recruited at CHU Sainte-Justine with a history of Infantile Myofibromatosis (shown in red) used for the linkage analysis.*

3. Sample selection and DNA extraction from blood, paraffin, frozen tissue

For the somatic study, the Infantile Myofibromatosis tumors were collected from patients who underwent surgery at the CHU Ste-Justine. In all cases, the diagnosis was confirmed by the Pathology Department at CHU Ste-Justine.

Out of the 41 Infantile Myofibromatosis cases diagnosed at CHU Ste-Justine, only 29 cases were retrieved from the pathology department archives. Among these, only 5 frozen tissue samples were available, of which only one case contained both normal and tumour tissues, the remaining four only contained the tumoral part. On the other hand, 29 paraffin blocks were recovered from the archives, of which many embedded tissues were previously fixed in Bouin (degraded DNA). Of these samples, only one was used for FISH study whereas the rest were eliminated from the study. A summary of the material available for the study is illustrated in figure 35.

It is evident that the study of orphan tumors has many challenges. A system has to be implemented to allow more efficient collection and maintenance of both normal and tumoral tissues to carry out more efficient molecular studies in the near future.

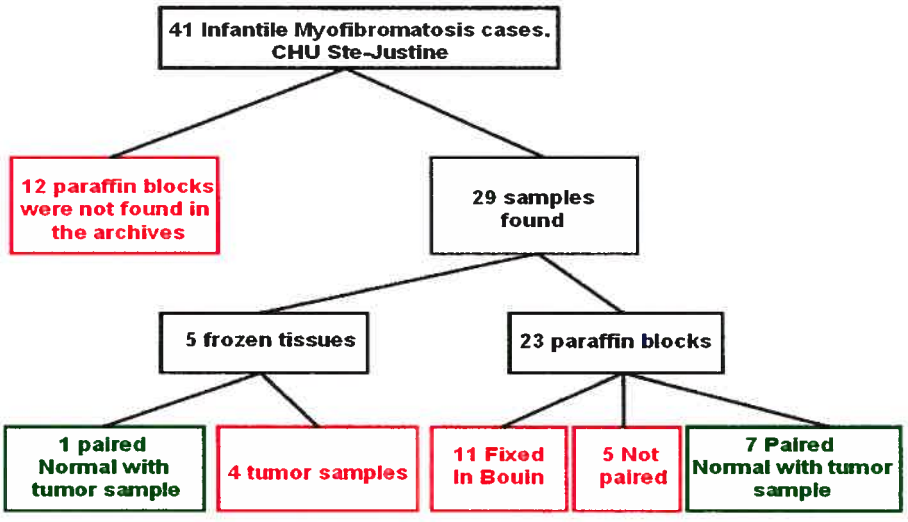
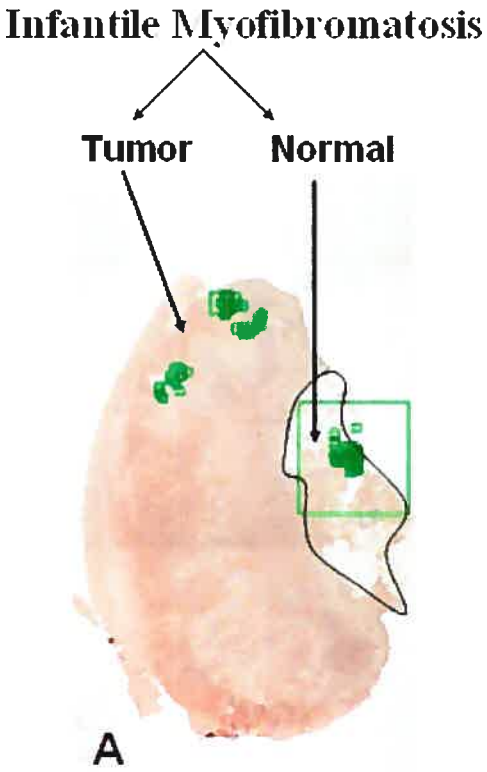


Figure 35: *Description of the material available to study Infantile Myofibromatosis. Red boxes indicate samples that were excluded from the LOH study.*

DNA samples from tumoral tissue and a corresponding normal tissue (blood or peripheral tissue) were isolated. On the other hand, limited frozen tissues were available for the study. DNA was extracted from all frozen tissues with standard methods using proteinase-K digestion and phenol/chloroform purification using ethanol precipitation. DNA collected was diluted in ultra pure, distilled, DNAse free and RNAse free water (Invitrogen Lot 1321975) and stored at -20°C. DNA from paraffin embedded tissues was isolated by phenol/chloroform purification, extract-N-Amp tissues PCR kit (Sigma , Lot 074K6043) or by Pico pure DNA extraction kit (Arcturus catalogue number KIT0103) after microdissection by AutoPixTM laser capture miscrodissection system (Arcturus) illustrated in figure 36.

For microdissection, Infantile Myofibromatosis cells from one patient were identified morphologically on hematoxylin and eosin (H&E) stained slides. Paraffin embedded tissue sections of the Infantile Myofibromatosis were deparaffinized, dehydrated and stained with H&E. Laser capture microdissection was performed according to manufacturer instructions. DNA was extracted in 10 µl of proteinase K extraction buffer (DNA Pico pure extraction kit, Arcturus Engineering, Mountain View, CA) and incubated for 20 hours at 65°C, followed by 5 minutes incubation at 95°C to inactivate proteinase K.



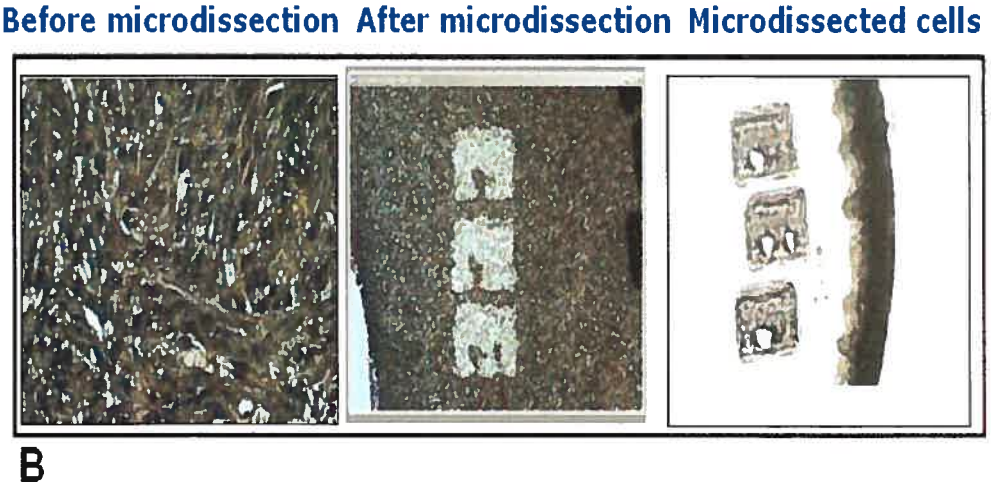


***Figure 36**: AutoPixTM laser capture miscrodissection system Arcturus.*

A. Infantile Myofibromatosis sample 04-7883 normal and tumoral cells during microdissection; green squares found on the slides demonstrate areas microdissected from the normal and tumoral tissues.

B. Laser-assisted microdissection was performed to collect cells

4. LOH analyses of chromosome regions 6q and 9q

Unfortunately, not all the diagnosed IMF samples could be retrieved. Also, DNA amplification with the microsatellite markers of chromosomes 6 and 9 was not a success when DNA was isolated from paraffin embedded tissues.

- DNA was never recovered from paraffin embedded tissue by phenol/chloroform.
- DNA isolation with the Sigma extract-N-Amp kit was not a success either since most of the time, both normal and tumoral DNA quality and quantity were not good enough for PCR amplification.
- Microdissection is a powerful tool. However, each DNA isolation step gives 10 ng (1ng/ µl) of DNA out of which 5 ng is used for one microsatellite marker amplification. Only two microsatellite markers were amplified for one patient due to time limitation.

Due to limited access to frozen tissues, only one paired normal and Infantile Myofibromatosis sample was analysed. Loss of heterozygosity was performed as discussed previously. The markers were selected from the Ensembl website and covered regions of chromosome 6q12-q15 and chromosome 9q. PCR amplification using frozen tissue was carried with 25 ng of genomic DNA template. The PCR cycles were adapted for each markers, including 10-minute denaturing step at 95°C, 26-30 cycles of 20-second denaturation at 95°C, 20-second annealing step at 54-59°C, 30-second elongation step at 72°C and 3-minute final extension step at 72°C PCR conditions using isolated genomic DNA and amplification with the Sigma kit was followed as per recommended by the manufacture. The PCR products were mixed with bromphenol blue containing loading buffer (LI-COR part # 830-05629), separated by electrophoresis on 6.5% polyacrylamide gel and detected by laser fluorescence using an automated gene sequencer (LI-COR).

Only one frozen tissue (sample 04-7783) with paired normal and tumoral sections was available. DNA was also extracted from paraffin embedded tissue for the same patient. Microsatellite markers amplification with DNA from frozen tissue permitted a higher throughput analysis compared with DNA extracted from paraffin blocks. Microsatellite analysis for Infantile Myofibromatosis sample 04-7883 using markers on chromosome 6 and 9 results are shown in table VIII; including the location of the markers and the matching results for each marker. No LOH was detected in this sample. Both DNA from frozen and paraffin tissues were used for this study.

***Table VIII**: LOH results using a variety of chromosome 6 and 9 microsatellite markers in patient's genomic DNA recovered from frozen tissue or paraffin embedded tissues.*

Patient 04-7883: DNA recovered from Frozen Tissue

Markers	Location	Result
D6S257	6p12.1	Retention
D6S430, D6S1694	6q12	Retention
D6S1718, D6S421, D6S1596	6q13	Retention
D6S456	6q14.1	Retention
D6S1627, D6S460	6q14.3	Retention
D6S1613, D6S462	6q15	Non-informative
D6S300	6q16.1	Retention
D9S273	9p21.11	Non-informative
D9S175	9p21.13	Retention
D9S167	9p21.32	Retention
D9S1812	9p21.33	Retention
D9S287, D9S1689	9p22.32	Retention
D9S261, D9S1677	9p31.3	Retention
D9S289	9p32	Retention
D9S1682	9p33.2	Retention
D9S290	9p34.11	Non-informative
D9S1838	9p34.3	Retention

Patient 04-7883: DNA recovered from paraffin embedded tissue

Markers	Location	Result
D6S1596	6q13	Retention
D9S289	9p32	Retention

5. FISH optimisation for paraffin embedded tissues

Since frozen tissues for Infantile Myofibromatosis were not readily accessible, we studied tissues embedded in paraffin by performing FISH. The objective of FISH analysis was also to determine the percentage of chromosomal deletions in patients with Infantile Myofibromatosis on chromosomes 6q and 9q. To optimize the methods on paraffin embedded tissues, we have preformed a series of experiments in which we studied the influence of fixation, deparaffinization, pre-treatment with proteolytic enzymes, and post hybridization conditions for FISH performed on paraffin-embedded and frozen tissues. This section describes:

a. Nucleus extraction from Infantile Myofibromatosis samples,
b. Hybridization specificity of the BAC probes,
c. BAC probes signal on Infantile Myofibromatosis tissues embedded in paraffin,
d. Pre-treatment to optimize BAC probe signals,
e. Transition for BAC probes to commercial probes ,
f. Results with Vysis Inc. commercial probes.

a. Nucleus extraction from Infantile Myofibromatosis samples

Tissue Core Collection

According to the diameter of the piece of tissue studied, 2 to 20 µm sections of paraffin-embedded tissue were produced on a microtome and collected in a centrifuge tube. For frozen tissues, touch preparation was done on slides.

Extraction of Nuclei

The paraffin was dissolved at room temperature with two 10-minute changes of xylene (10 ml each) in the centrifuge tube. The tissue was then rehydrated with 10 ml of 90%, 70%, and 50% ethanol (EtOH) for 5 minutes each. The 50% EtOH was removed and replaced with 10 ml 0.9% NaCl for 5 min. The tissues were transferred to a microtube and enzymatic digestion was then performed by adding various volumes of freshly prepared proteinase K solution (5 mg proteinase K (Invitrogen), 50 µl 1M Tris-HCl (pH 7.5), 20 µl 0.5 M EDTA (pH 7.0), 2 µl 5 M NaCl, make up to 1 ml with filtered double distilled water) to the microcentrifuge tube. The specimen was incubated at 37°C for 30 minutes to 2 hours. To aid with enzymatic digestion, the sample was vortexed during the incubation period. The purification of the released nuclei was done by transferring the digested fluid onto a 60 µm nylon mesh (Millipore NY6004700) and nuclei remaining in the mesh were washed out by 4 ml 1XPBS, passed through the mesh and collected in the 15 ml plastic tube. The extracted nuclei were recovered by centrifugation and resuspended in 300 µl of 1XPBS. The suspension of 20 µl was distributed on clean dry slides and was left to dry overnight. After washing the slides with formalin buffer to remove crystallized PBS salts, an evaluation of the success of the nuclear extraction was performed by phase contrast light microscopy.

Pre-treatment and FISH

Slide preparations were circled to define the hybridization site and mark the location where FISH probes were applied. As in conventional FISH approach, a pre-treatment of the slides with RNase and pepsin followed by a postfixation with formalin-buffer is required to reduce the background. In brief, 100 µl of RNase solution covering the surface the extracted nuclei are incubated in a humid

chamber for 15 minutes at 37°C. After washing the RNase solution, slides were transferred to a Coplin jar containing freshly prepared pepsin solution. The slides were then dehydrated with 70%, 85%, and 100% EtOH for 2 minutes each.

FISH and Post-hybridization Wash

For the optimisation of FISH protocol, various volumes of the BAC probes were used to increase the hybridization signal. Also, different modifications in the concentration of solutions, denaturation time and denaturation temperature were adjusted each time. Probes, provided by the TCAG genome resources facility (The hospital for Sick Children), were isolated, digested and labelled SpectrumGreen or SpectrumOrange and finally re-suspend in 20 µl DDH2O by the manufacturer. The description of the BACs used in this study is listed in table IX. The BAC probes were mixed with 10 µl of hybridization buffer (Vysis, Inc.) and with 1 µl Human Cot 1 (Invitrogen Cat. 15279-011) before hybridization. The Cot-1 DNA binds to the probes repeated sequences which in turn cannot hybridize to the chromosomal repetitive DNA leaving the unique sequences contained in the probe single-stranded and free to hybridize to the chromosomal DNA. The probe ordered from Vysis Inc. was not mixed with Human Cot 1. Two types of denaturation procedures were used. For metaphase chromosomes, probes were denatured in a 75°C water bath and the slides with the chromosomes were immersed in 70% formamide. The denatured probe solution prepared was applied to the hybridization site and covered with a coverslip. For interphase nuclei isolated from both frozen and paraffin embedded tissues, probe solution was applied to the hybridization site and covered with a coverslip followed by denaturation on a hot plate. The use of a hot plate seemed to denature chromosomes more efficiently compared to manually denaturation using water bath. This is probably due to a better control of the temperature and a more efficient contact on a hot plate compared to the water bath. Specimens were stored in a humide chamber at 37°C for 48 hours. After hybridization, coverslips were gently removed and the slides were washed in a Coplin jar filled with 50% formamide/ 4XSSC followed by 4x SSC /0.1% Tween-20. The washing step was kept minimal for retention of hybridization probe signal. Nuclei were counterstained with a mixture of 10 µl of DAPI II staining: 125ng/ml (VYSIS)

(antifade is already added in the mixture). A 24 x 50-mm coverslip was then placed over the hybridization sites.

Microscopy

Analysis was done using a fluorescence microscope equipped with a 100-W mercury lamp. To view signals a dual-pass SpectrumOrange/SpectrumGreen filter set was used. Individual fluors were observed using a single-pass SpectrumGreen or single pass SpectrumOrange. A total of 100 up to 200 consecutive qualifying interphase nuclei were scored for each specimen. Since there is a possibility that the nucleus is cut leaving out genomic material, it was important to score each probe signal individually. The use of control probe is necessary to validate chromosome deletion that could be present in nuclei.

***Table IX**: Summary of DNA FISH Probes Used in this Investigation*

Probes	Hybridization loci	Used to detect	Probe Size
RP1-223E3	6q14.2	del(6)(q12q15)	13 Kb
RP1-120N9	6q14.2	del(6)(q12q15)	170 Kb
RP11-427L11	9q31.3	Monosomy 9q	195 Kb
RP11-553M22	16q22.1	Trisomie 16q	200 Kb
BCR/ABL ES Dual Color Translocation Probe	9q34, 22q11.2	Monosomy 9q	9q34=650Kb 22q11.2=300Kb

b. Hybridization specificity of the BAC probes

Some clones from the RPCI-11 library can hybridize to more than one chromosomal location [99]. For this reason, the BAC probes were tested against normal metaphase chromosomes obtained from the department of pathology to confirm the probe location and signal number illustrated in figure 37). All three probes, RP1-120N9, RP11-427L11, RP11-553M22 showed a high specificity for their respective chromosome except RP1-223E3. The latter was therefore eliminated from the study.

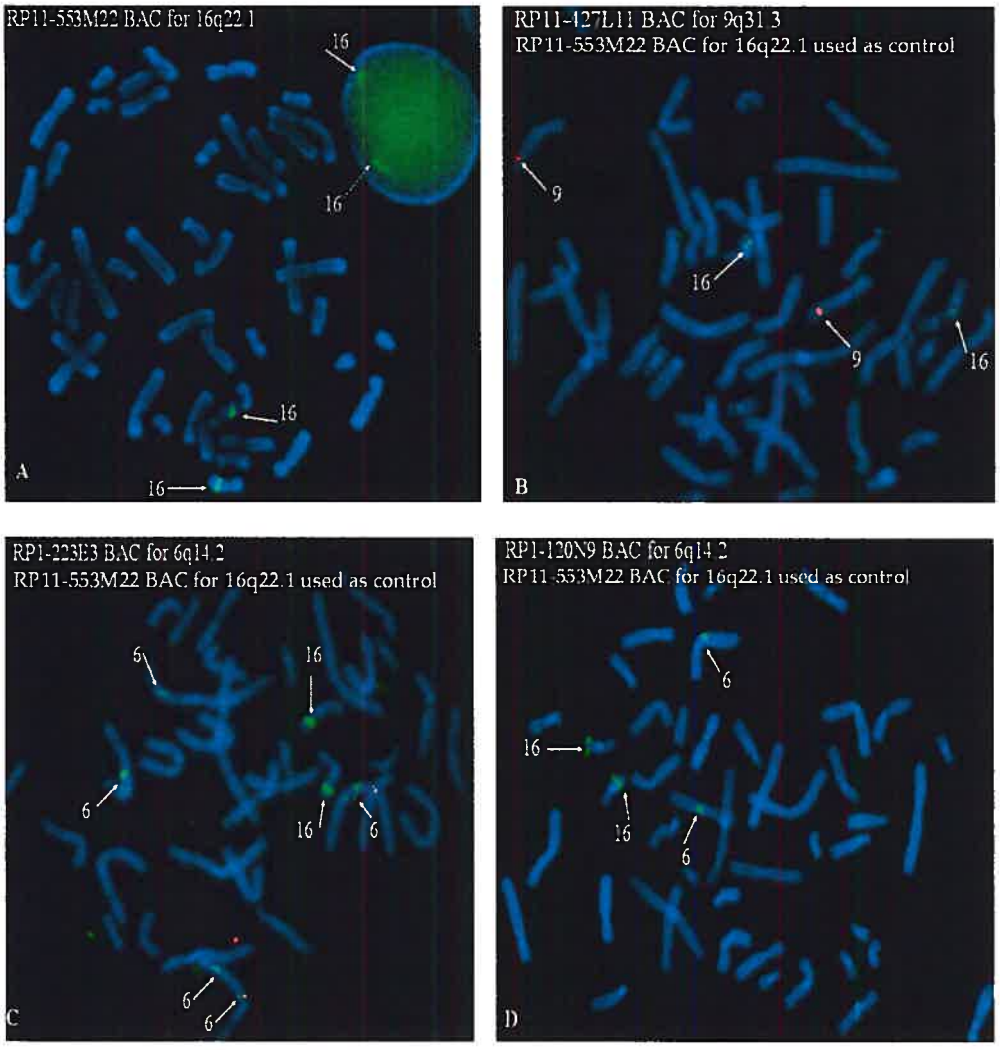


***Figure 37**: BAC probes test for their hybridization specificity.*

A) RP11-553M22 BAC for chromosome 16 used as control for B,C,D, B) RP11-427L11 BAC for chromosome 9, C) RP1-223E3 BAC and D) RP1-120N8 chromosome 6.

c. BAC probes signal on Infantile Myofibromatosis tissues embedded in paraffin

To test if these signals were detectable on paraffin embedded tissues, nuclei were extracted from formalin fixed/paraffin embedded tonsil and interphase FISH was carried out with the BAC probes. All three, RP11-553M22, RP11-427L11, RP1-120N8 BACs, hybridized well on paraffin embedded tissues and the signal was easily detectable with the fluorescence microscope (figure 38). However, for the Infantile myofibromatosis extracted nuclei, the only probe detectable with the fluorescent microscope was RP11-553M22 (figure 39).

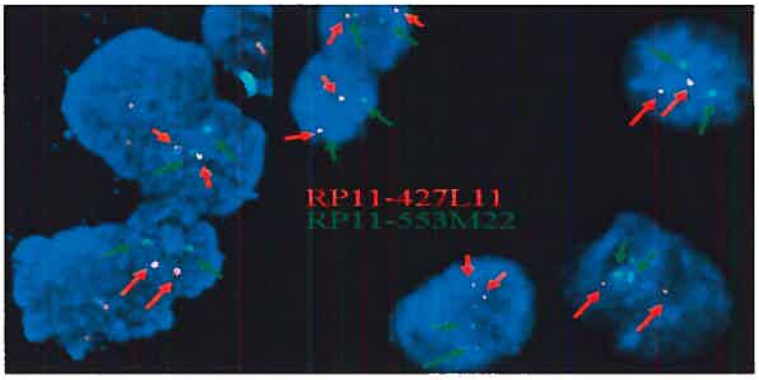


***Figure 38**: Tonsil tissue embedded in paraffin was used to test the BAC probes. After nucleus extraction, interphase FISH was performed with BAC RP11-427L11 shown in red and BAC RP11-553M22 shown in green.*

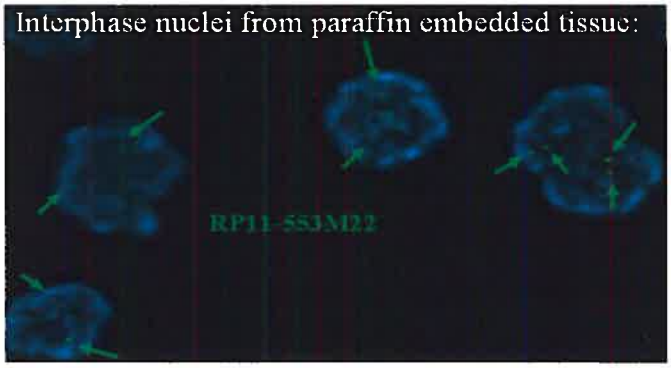


***Figure 39**: BAC Probe signal MFI paraffin embedded tissues. The only BAC signal detectable on Infantile Myofibromatosis tissues was BAC RP11-553M22 shown in green.*

d. Pre-treatment to optimize BAC probe signals

Since the hybridization signal with the BAC probes detection was low on the Infantile Myofibromatosis interphase nuclei, additional pre-treatment was done on the slides. Recently, Paternoster et *al* [100] described a successful pre-treatment of the nuclei extracted form paraffin-embedded tissues to allow better demasking and denaturation of the target DNA prior to FISH. They recommended boiling the sections in citric or sodium bisulfate buffer prior to FISH. Sodium bisulfate pre-treatment ameliorated the condition of the BAC probe hybridization on Infantile Myofibromatosis interphase nuclei compared to the citrate buffer treatment or no treatment at all. However, the probe signals were not intense enough for analysis with the naked eye, the probe signals were only detectable by taking pictures with the fluorescence microscope of individual cells (figure 40).

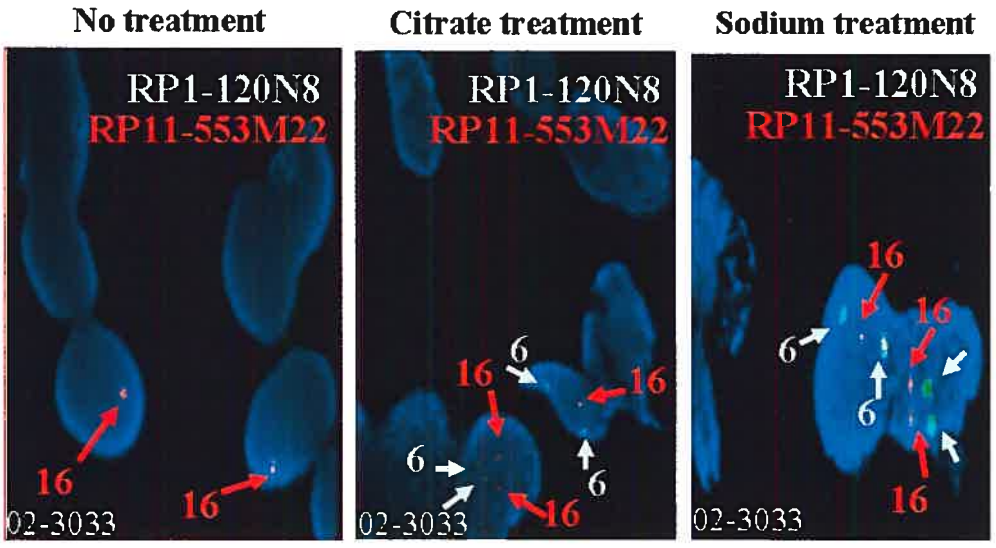


***Figure 40**: Interphase nuclei form Infantile Myofibromatosis were pre-treated with citrate and bisulfate buffers. The signal obtained after hybridization of the BAC probes was best when the nuclei were treated with sodium bisulfate buffer.*

e. Transition from BAC probes to commercial probes

Due to the weak signal of the BAC probes even after pre-treatment optimization, LSI bcr/abl ES (extra signal) dual color DNA probe was purchased from Vysis Inc. This probe normally hybridizes to chromosome 22q11.2 (breakpoint cluster region SpectrumGreen) and to chromosome 9q34 (abl oncogene SpectrumOrange) to detect the t(9;22)(q34;q11.2) which defines the Philadelphia chromosome (figure 41). In our study, this probe was used to detect the percentage of chromosomal deletions of chromosome 9q in Infantile Myofibromatosis cases and chromosome 22 was used as a hybridization control on interphase nuclei.

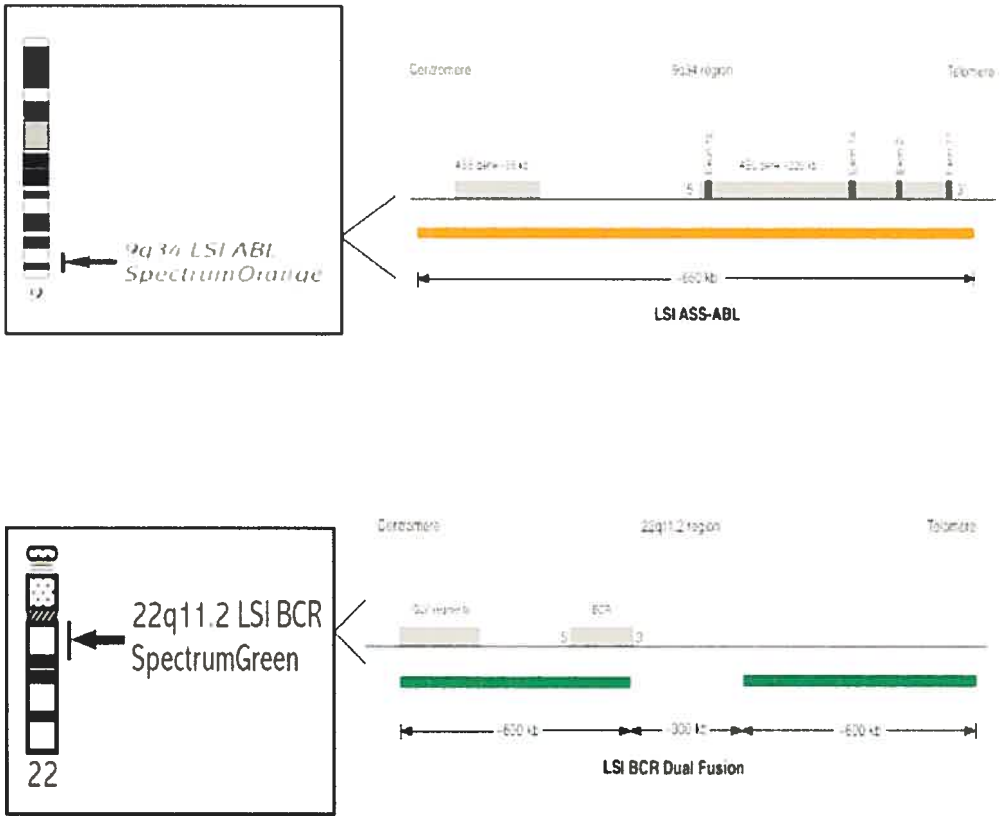


***Figure 41**: LSI bcr/abl ES (extra signal) dual color DNA probe. This probe hybridizes to chromosome 22q11.2 (breakpoint cluster region SpectrumGreen) and to chromosome 9q34 (abl oncogene SpectrumOrange). The probe was used to detect the percentage of chromosome 9q deletion in Infantile Myofibromatosis.*

f. Results with Vysis Inc. commercial probes.

In total, 9 Infantile Myofibromatosis samples embedded in paraffin were used to determine the percentage of possible chromosome 9 deletion. A total of 100 up to 200 consecutive qualifying interphase nuclei were scored for each patient. Frozen samples were not examined since all gDNA was isolated for genotyping with the SNP array.

Interphase nuclei for sample CH-00-7939 could not be analysed after probe hybridization. Due to the high denaturing temperature, the nuclei lost their morphology and probe intensity was not readable. CH-85-7297 was fixed in Bouin and nuclei isolation was impossible. For the remaining samples: CH-00-7939, CH-03-3055 and 03-701, an insufficient number of nuclei was isolated form the blocks. Additional nuclei purification is needed in those cases.

Signal intensity was quantified for both target chromosome 9q region and chromosome 22, used as a control. Since sections of the paraffin blocks were used for nuclei isolation, it was important to differentiate loss of probe signal due to nuclei section cuts after purification or due to true loss of chromosome regions. After examining all the paraffin samples used, we noticed an approximate equal loss of target (chromosome 9q) and control (chromosome 22) signal indicating that some sections of the nuclei were cut during the purification step. Results are shown in table X. Chromosome 9q deletion was not observed in any of the IMF samples analysed with the LSI bcr/abl ES dual color DNA probe; loss of signal was due to partial loss of nuclei during purification step.

***Table X**: Results of FISH on interphase nuclei using the LSI bcr/abl ES dual color DNA probe to detect chromosome 9 deletions in Infantile Myofibromastosis. The first column indicates the samples analysed by FISH, the second column indicates the number of nuclei with retention of both chromosome 9 and 22 signals, the third column indicates the number of nuclei with retention of chromosome 9 signal thus indicating a loss of chromosome 22 signal, the fourth column indicates the number of nuclei with retention of chromosome 22 signal thus indicating a loss of chromosome 9 signal.*

Sample Number	spectrum orange (chr9) and green (chr22)	spectrum orange (chr9) only	spectrum green (chr22) only
Paraffin tissue:			
CH-02-3033	200	7	8
CH-00-8613	200	12	11
CH-01-4872	100	11	10
CH-01-2717	200	13	13
CH-00-7939	NA	NA	NA
CH-85-7297	NA	NA	NA

6. Infantile Myofibromatosis SNP array

Since no chromosome deletion was detected by FISH analysis, the gDNA extracted from the frozen fragment of three Infantile Myofibromatosis was analysed for chromosomal aberration and on the Sentrix Human-1 SNP BeadChip. Due to patient consent limitation, only three out of the six infantile myofibromatosis frozen tissue samples were analysed by SNP array. The tumoral gDNA was compared to the reference cluster for all IMF cases; no normal samples were used for the analysis.

No chromosomal aberrations were observed throughout the whole genome of the three samples analyzed (figure 42, 43, 44). Results of chromosome 6 and 9 for each patient are illustrated by the images bellow.

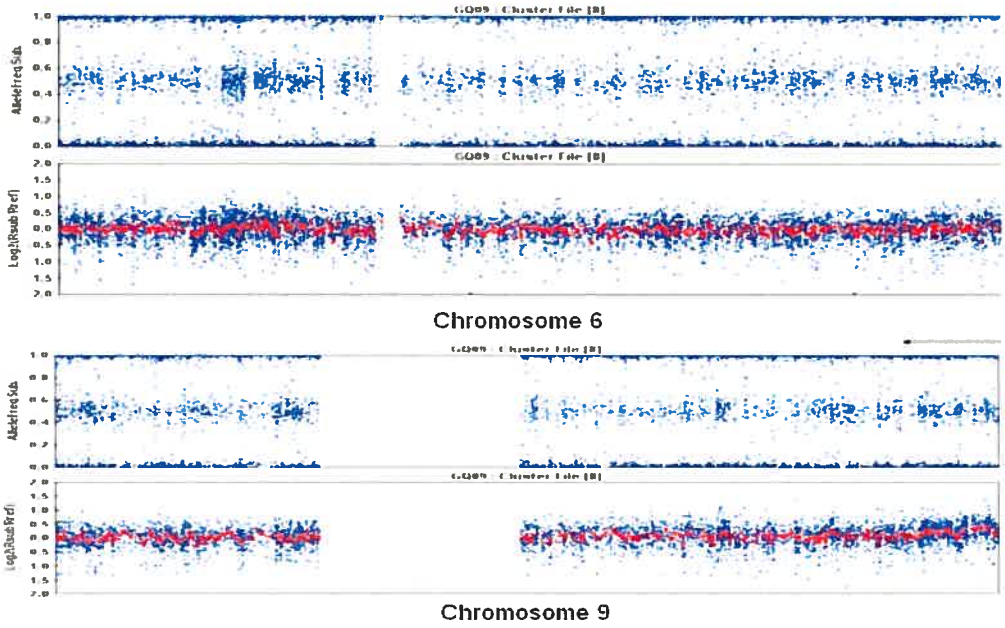


***Figure 42**: Infantile Myofibromatosis sample 04-7883 DNA analysis with SNP array*

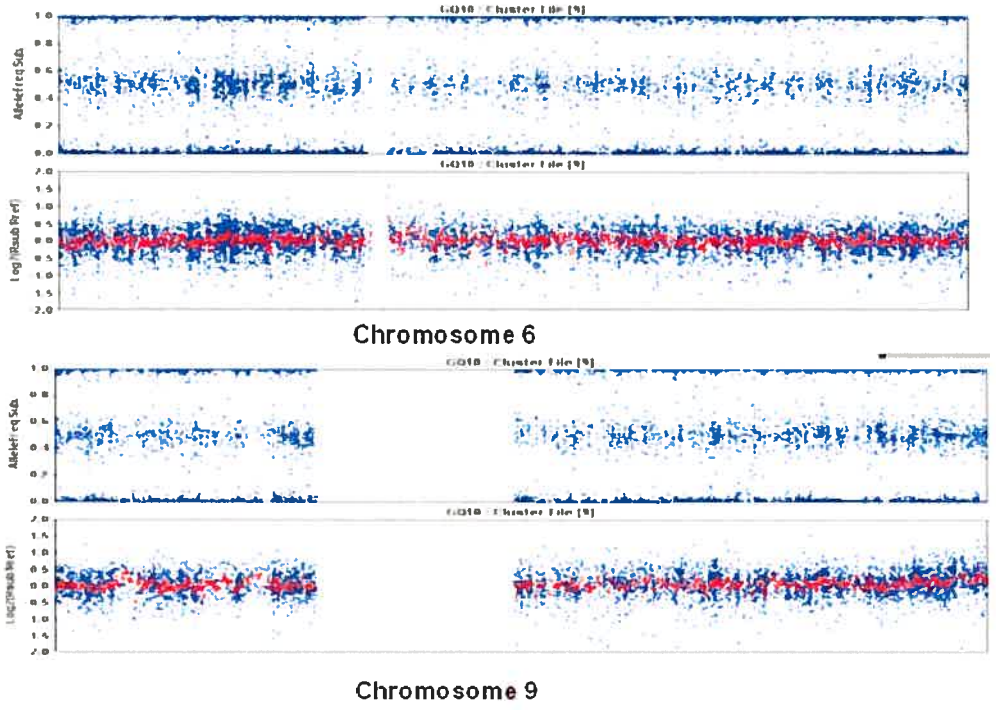


***Figure 43**: Infantile Myofibromatosis sample 2 DNA analysis with SNP array*

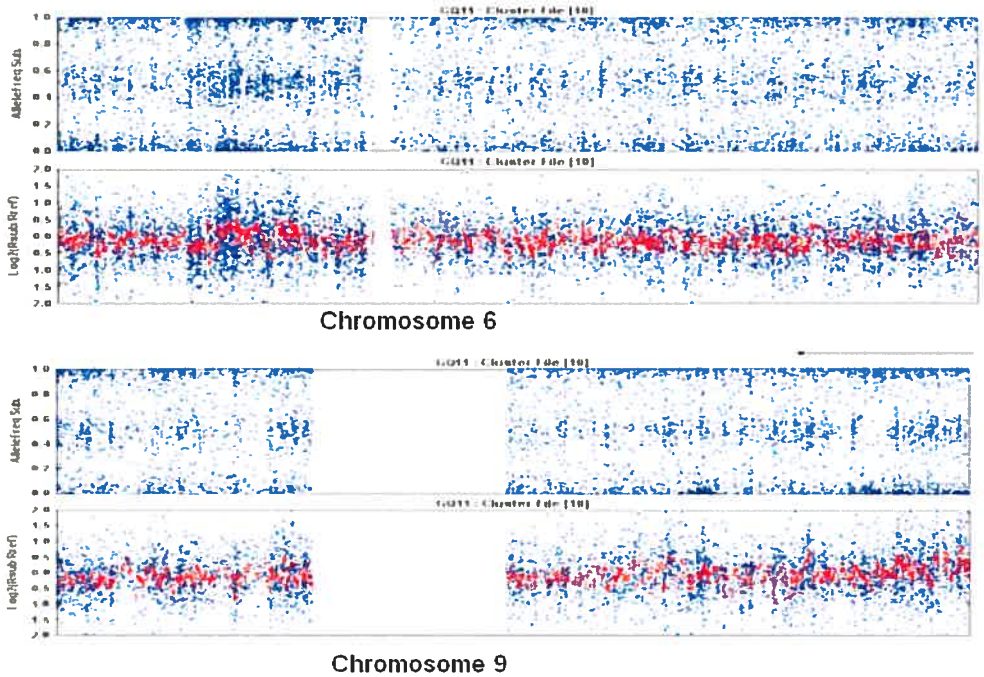


***Figure 44**: Infantile Myofibromatosis sample 3 DNA analysis with SNP array. Even with poor DNA quality, Sentrix Human-1 SNP BeadChip is capable to perform tumoral genotyping.*

C. Rational for Intestinal Epstein-Barr virus-associated smooth muscle

We thought it would be interesting to study a second rare paediatric tumour encountered in a patient treated at CHU Ste-Justine. We choose this tumour since there is a limited amount of molecular data available in the literature regarding this entity and the evaluation of the karyotype with metaphase chromosome was not successful.

Techniques used in this case were:

- Somatic study was carried out by
    - HUMARA assay to prove the clonality of the tumour
    - FISH on interphase nuclei
    - Sentrix Human-1 SNP BeadChip array

1. Introduction

Epstein-Barr virus (EBV) is a human herpes virus whose oncogenic expression varies according to the immune status of the host. In immunocompromised patients, EBV-associated tumors consist of post-transplantation lymphoproliferative disorders (PTLD), EBV-associated lymphomas and immunodeficiency-related smooth muscle neoplasms.

Smooth muscle tumors (SMT) rarely occur in the paediatric population, but an increased incidence of these neoplasms is observed in the setting of immune deficiency, whether congenital or acquired, in which cases they are associated to the EBV. These EBV-related SMTs arise with increasing frequency following iatrogenic immunosuppression consecutive to solid organ or bone marrow transplantation. Recently, 19 cases of EBV-associated smooth muscle tumors (EBV-SMT) were reviewed, of which 14 had occurred in children [101]. Only one karyotype was published in the literature from a 10-year-old patient with an EBV-related liver tumour occurring after kidney transplantation. Cytogenetic analysis revealed a single clonal cell population showing [102]:

46,XY,del(2)(p23),der(3)t(2;3)(p23;q29),der(21)t(Y;21)(q12;p13)

No other karyotypes are available and the clonality of the tumour was never previously examined.

2. Sample selection and DNA extraction

A 14-year-old girl underwent living-related kidney transplantation due to end-stage renal failure at CHU Ste-Justine. Primary EBV infection was detected two weeks after surgery. Multiple synchronous abdominal lesions were surgically removed. With the patient parent's consent, genomic DNA was extracted from a frozen fragment of a smooth muscle tumour lymph node metastasis with standard methods using proteinase-K digestion and phenol/chloroform purification. DNA collected was diluted in ultra pure, distilled, DNAse free and RNAse free water (Invitrogen Lot 1321975) and stored at - 20°C.

3. Clonality assessment by HUMARA analysis

Tumors are formed by the clonal expansion of a single precursor cell that has incurred the genetic damage and therefore, they are referred to as being monoclonal [103]. To assess whether the Epstein-Barr virus-associated smooth muscle was effectively derived from a single precursor cell, analysis of X-chromosome inactivation patterns was carried out. The human androgen receptor (HUMARA) locus is especially useful for clonality studies [104]. To our knowledge, this information was not previously published in the context of an Epstein-Barr virus-associated smooth muscle.

The genomic DNA previously extracted from the frozen fragment of a smooth muscle tumour lymph node metastasis was used to test the clonality of the tumour. PCR amplification of the polymorphic CAG repeat at the HUMARA locus was performed in tandem on undigested and on HhaII digested DNA as described by Allen et al [105]. An M13 universal primer sequence tail covalently attached to the 5' terminus of the forward amplification primer facilitated electrophoretic analysis of the PCR products using infrared detection (DNA Analysis System from LI-COR Biosciences). The PCR results demonstrate that the EBV virus-associated smooth muscle tumour derived from a single precursor cell (figure 45.)

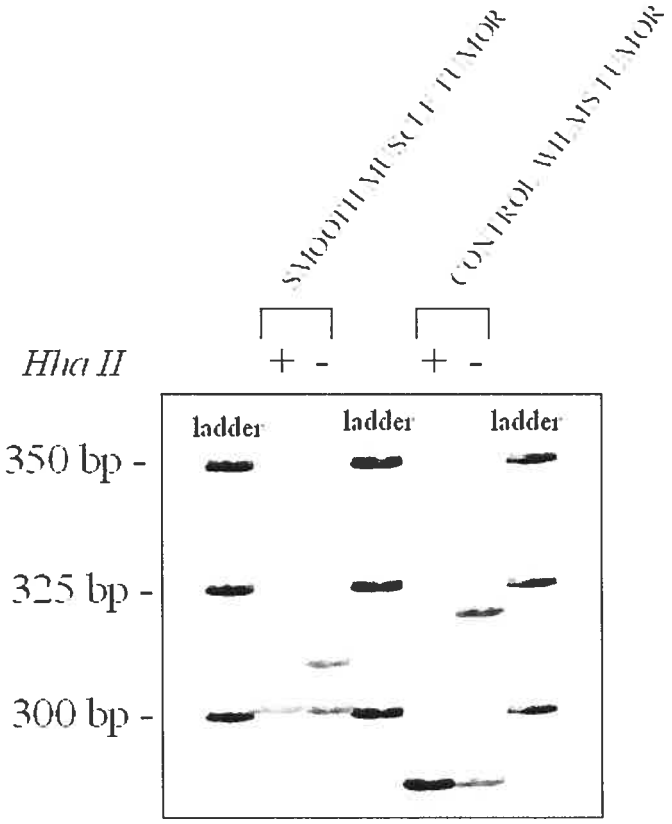


***Figure 45**: HUMARA results: prior to the PCR, the DNA of the smooth muscle tumour was either digested or not digested with the methylation-sensitive enzyme HhaII which cleaves the restriction sites on the active (nonmethylated) chromosome. The PCR does not then amplify these active alleles in monoclonal tissues. DNA extracted from monoclonal Wilms' was used as a control.*

4. FISH

Possible 2p23 chromosome deletion was examined in the EBV-SMT. FISH was done with a freshly made imprint touch preparation obtained from the tumour tissue. Interphase nuclei were hybridized with labelled DNA probes LSI ALK dual color break apart rearrangement probe that mapped to *ALK* (2p23) (illustrated in figure 46). By FISH analysis, we were able to demonstrate a homozygous deletion (loss of both probe signals in one nucleus) of the chromosome 2p23 in about 80% of the tumour cells, involving the *ALK* gene located on 2p23. To verify if the whole chromosome 2 was lost, interphase nuclei were hybridized to CEP 2 (D2Z1) probe which targets the centromere of chromosome 2. There was no loss of the chromosome 2 centromere, further confirming that the homozygous deletion is limited to an incomplete portion of the short arm of chromosome 2. The results are illustrated in figure 47.

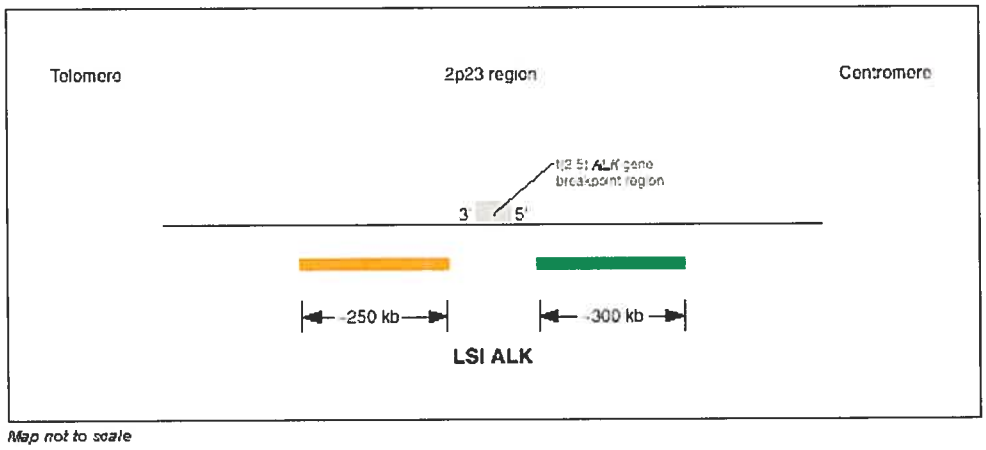
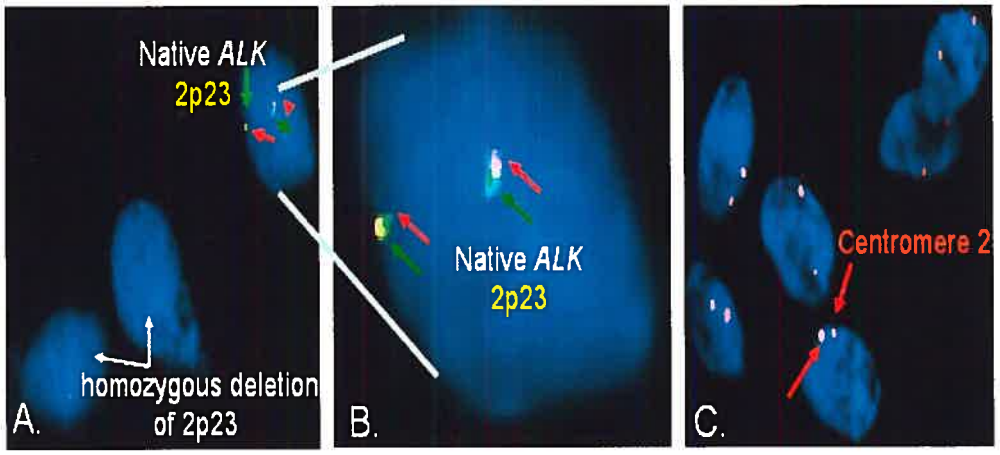


***Figure 46**: LSI ALK Dual Color, Break Apart Rearrangement Probe. When hybridized with the probe, the 2p23 ALK region in its native state will be seen as two immediately adjacent or fused orange/green (yellow) signals*



DNA probes for *ALK* (2p23) showing loss of both dual color signal in 80% of nuclei

DNA probes for *ALK* (2p23) showing dual color

Probe detecting centromere of chromosome 2

***Figure 47**: FISH results on interphase nuclei from EBV-smooth muscle tumour.*

***A.** Homozygous deletion of chromosome region 2p23 in 80% of the EBV-SMT nuclei.*

***B.** Retention of chromosome 2p23 probe signal in 20% of the normal EBV-SMT nuclei.*

***C.** Probe signals showing the retention of centromere on both chromosomes arm in 100% of EBV-SMT nuclei.*

5. SNP array result

We wanted to know if the results obtained by FISH were reproducible with the SNP array. The genomic DNA was extracted from a lymph node containing the smooth muscle tumour metastasis. Unfortunately, the 2p23 deletion was not detected with the Illumina SNP array (figure 48). Normal gDNA contamination would be a possible explanation for the discordance between FISH and SNP array results. The tumour allele frequency could be indistinguishable from the normal allele frequency; the normal heterozygote allele frequency could in this case mask a homozygous deletion.

This case illustrates a possible limitation of the SNP array. Further investigations are needed to support this argument since the two methods (FISH and SNP array results) are contradictory. A second FISH analysis is necessary with the addition of an internal control to verify proper hybridization of the dual color break apart rearrangement probe. Also, a second genotyping of the EBV-SMT gDNA could be necessary, performing a dissection of the tumoral tissue from the normal tissue to achieve ultimate gDNA purity. Finally, if the tumour contains a translocation, it can not be detected by the SNP array.

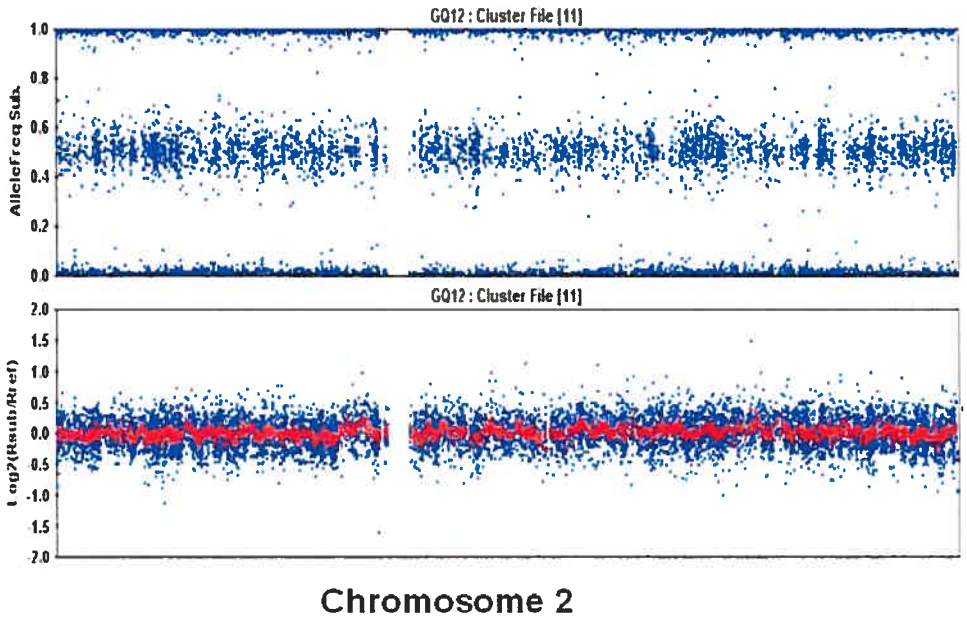


***Figure 48**: EBV-SMT genotyping result, showing chromosome 2t. Unlike with the FISH analysis, no deletion of the chromosome 2p23 in the intestinal Epstein-Barr virus-associated smooth muscle was observed by SNP array.*

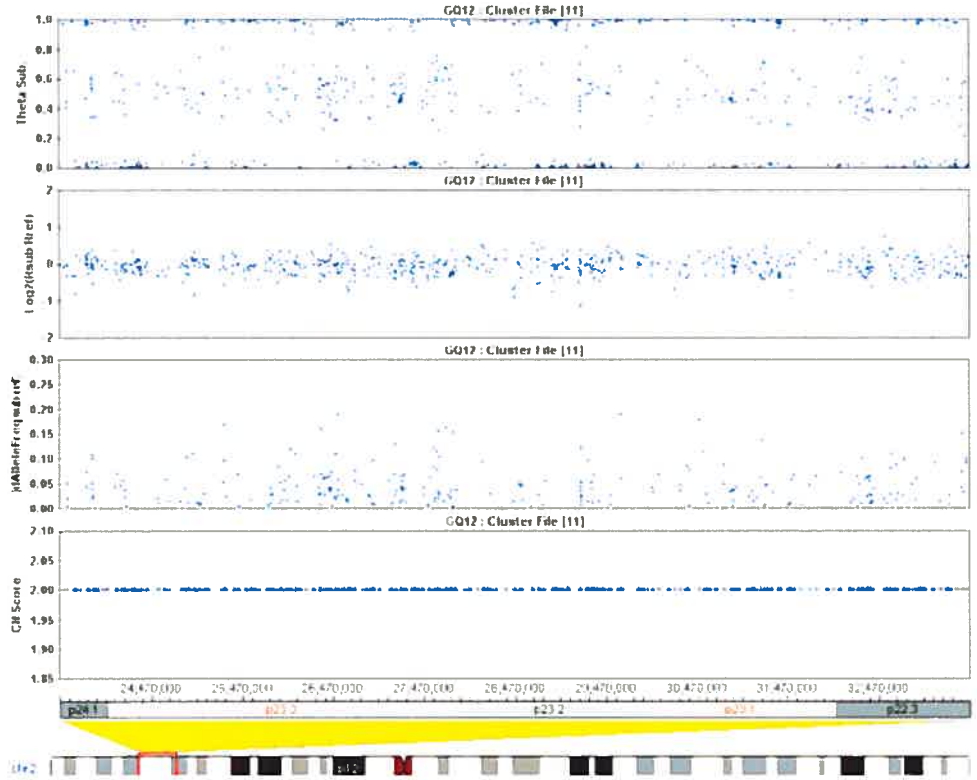


***Figure 49**: Zoom in image of chromosome region 2p23 without any detection of a deletion.*

CHAPTER V

- Discussion -

We have explored a variety of technologies to examine chromosomal abnormalities in rare paediatric cancers including microsatellite marker amplification with DNA from frozen tissue and paraffin embedded tissue, FISH with BAC and commercial probes and finally, high through put SNP array genotyping. There are many limitations in conventional techniques but with the use of high throughput arrays, analysis of rare paediatric cancers is made easier, more productive and efficient. Limitations and advantages of the different techniques employed in this study are illustrated bellow.

1. Microsatellite amplification for LOH studies

1.1. Frozen tissue

PCR amplifications for LOH studies is very productive when frozen tissue is used. However, a major limitation in this study was the limited access to both normal and tumoral samples. A system has to be implemented to allow more efficient collection and maintenance of both normal and tumoral tissues to carry out molecular studies in the near future. Also, analysis of LOH studies can be misleading and needs to be repeated when uncertainties arise which is cost and time consuming. These uncertainties can arise due to the different definitions for LOH thresholds used; most studies take a threshold of 50% reduction to be significant, although some use less-stringent arbitrary thresholds {Tomlinson, 2002 #48}.

1.2. Paraffin embedded tissues

Paraffin embedded tissues are like treasures in the field of Pathology. They can be kept for a long period of time and serve for retrospective studies. However, the DNA is mostly degraded in fact more than half of the recovered paraffin blocks for this study were embedded in Bouin. Also, paraffin blocks often contain a mixture of normal and tumoral tissues. For these reasons, microdissection is highly recommended since is allows to recover a larger amount of DNA than with normal chloroform/phenol. It also allows a perfect way to separate different types of tissues (normal VS tumoral). However, it is highly time consuming. One can spend a whole day microdissecting one slide to recover the necessary amount of

tissues for a PCR reaction. Also, only 10 ng (1ng/ µl) of DNA is recovered each time microdissection is performed with Arcturus. Therefore, high throughput allelotyping was not feasible with this technique.

2. Karyotype tools application and limitations

2.1. Conventional karyotyping

Conventional karyotyping is a reliable and highly accurate means of diagnosis of a wide range of chromosome abnormalities, but it is really possible to establish primary tumors in tissue culture. In addition, cell culture is very expensive in terms of labour, reagents and media. But, it has an advantage over CGH/SNP-CGH arrays studies since balanced translocations are not detected by SNP-CGH.

2.2. FISH

One application of FISH involves the hybridization of probes to interphase cells. This is extremely beneficial when it is not possible to prepare metaphase spreads. FISH allows the screening a large number of cancer cells in search of an extremely rare subpopulation. Interphase cytogenetics also allows one to precisely define the cell pool carrying chromosomal abnormalities, to identify whether aberrant cells exist in clonal patches or as isolated events. Chromosome aberrations can be observed on a cell-to-cell basis rather than as a population. In addition, interphase FISH can be performed on paraffin-embedded, formalin-fixed tissue sections thereby allowing researchers to retrospectively analyze samples and correlate chromosome aberrations with biological and clinical endpoints. However, there are only a limited number of commercial probes available and BAC probes require optimizations not easily achieved by diagnostic laboratories due to time and budget constraints.

2.2 A- Advantages of FISH:

Comparative advantages of fluorescence in situ hybridisation (FISH) over loss of heterozygosity (LOH) analysis and SNP array as a method for detecting chromosomal aberrations include:

- DNA extraction and blood samples are not required.
- Can be performed on archived tissues and doesn't require metaphase spreads, but can not be used when tissue is fixed in Bouin.
- Can detect heterogeneity of cells with respect to specific deletions.
- More sensitive than LOH analysis and CGH in identifying deletions in a mixed population of cells (i.e. mixture of and normal parenchymal cells).
- In contrast to LOH analysis, FISH can distinguish deletion from homozygosity.
- Contaminating DNA is not a consideration.

  This point is particularly important to illustrate that the SNP array might not be able to detect homozygous deletion when gDNA contain both normal and tumoral DNA, as seen when EBV-SMT was examined with FISH and SNP array.

2.2. A- Main Disadvantages of FISH Techniques:

- Need specialised camera and image capture system.
- Limited number of commercial probes available, BAC probes need optimization
- Hybridization variation for different tissues. This makes the technique difficult to standardize across different laboratories.
- FISH is only suitable in cases with moderate to high numbers of abnormal cells. Cell selection technique is needed if malignant cells of interest are present at a low level.
- FISH will only provide information about the probe being tested at the specific locus, other aberrations will not be detected whereas SNP array gives a more global view of the genome.

Therefore, an easy, rapid, standardized and robust technology capable of identifying genome wide aberrations at ultrahigh resolution would represent an important advance in clinical diagnostics.

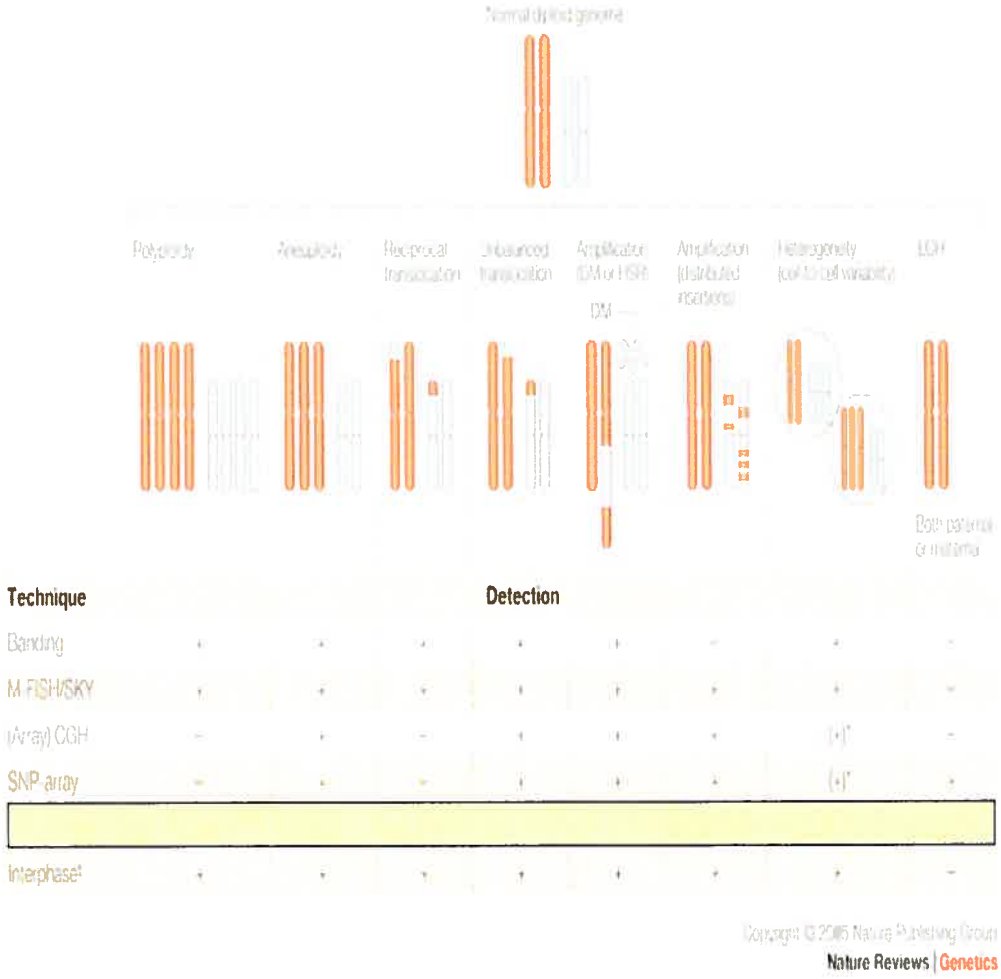


***Figure 50**: Comparison of cytogenetic techniques for identifying chromosomal abnormalities. Combination of FISH and SNP array covers a wide range of chromosome abnormalities [13].*

3. SNP array validation

3.1. Validation of the SNP array

We demonstrate the utility of SNP-CGH with the Infinium whole-genome genotyping BeadChips, assaying 109,000 SNP loci, to detect chromosomal aberrations in tumour samples at sub-100 kb effective resolution. Important aspects of the array were examined to lend validity and power to the method. Only recently, Illumina published their results on different aberrations observed in human cancer samples by using the same SNP platform [71]. All possible detectable aberrations published in the article were also observed in our Wilms' samples and included:

1. Heterozygote Deletions (LOH)
2. Amplifications
3. Duplication
4. Copy-neutral LOH.

We also illustrated that data obtained were reproducible with microsatellite marker amplifications to detect LOH. Important advantages of Illumina 109K Sentrix Human-1 SNP array include the possibility of analysing tumoral samples without comparison to its normal control. This is a tremendous advance in paediatric cancer research where access to both normal and tumoral samples is limited.

It is important however to keep into account that the LOH score algorithm used in the BeadStudio 2.0 LOH Plus module does not incorporate haplotype structures and assumes that heterozygote frequencies in the reference cluster are representative of the population under study. The CEPH panel of 120 samples include Caucasian, Han Chinese, Japenese and Yoruba HapMap populations. False positives in the LOH score may occur due to some SNPs being rare in the studies population while being diverse in the panel. Therefore, additional experiments might be necessary to help to confirm the validity of our genotyping data obtained from the analysed Wilms' samples. We still believe that our

preliminary studies along with the ones obtained during the recent validation (august 2006) of the SNP platform by Illumina [71, 72] demonstrate the possibility to use the platform for karyotyping numerous cancers.

3.2. Resolution of the Sentrix Human-1 SNP BeadStudio array

As the density on Illumina's BeadChips increases, the effective resolution for discovering ''needle in a haysatack'' genetic event will can be made possible. As mentioned previously, Illumina provides three types of bead arrays: the 109K Sentrix Human-1 SNP, the 317K HumanHap300, and the higher-density conventional or custom HumanHap550 (550k tag SNPs). Illumina is also advertising its future release of the HumanHap650Y BeadChip which offers the full SNP content of the HumanHap550 BeadChip plus an additional 100,000 tag SNPs selected from the Yoruba population. This product is ideally suited for whole-genome association studies for populations including individuals with African or with African-admixed genetic backgrounds but can also be used for LOH/DNA copy number studies. (http://www.illumina.com/products/snp/whole_genome_genotyping.ilmn). With the increase of the resolution of these arrays, tumoral genotyping will allow to find very small regional deletion or amplifications.

Illumina array's effective resolution is calculated by taking the SNP density into account along with a given window of SNPs. For example, the HumanHap550 has a median SNP spacing of 5kb. Using a 10SNP window, the effective resolution is 5kb times the 10 SNP window (~50kb). While there are regions that have many more (or less) SNPs, the 10 SNP window [71] was chosen since it provides the lowest false positive rate and includes the highest number of heterozygous SNPs based on an average heterozygosity of ~35%. For the Human-1, the median spacing is 26kb so the effective resolution is approximately 260kb. Therefore, the 100K SNP array could miss smaller chromosome aberrations. Considering the effective resolution of CGH varies from ~20 Mb with metaphase spreads down to ~100 kb with high-density array CGH using BAC or oligonucleotide arrays [71], the 100K SNP array does not provide a higher resolution for the detection of small chromosome aberrations. Using a greater

density array, in particular the 317K array, would allow the discovery of more small aberrations than the 109K.

3.3. Genomic quantity necessary to carry out genotyping

A down fall of Illumina array is the requirement of a high starting material (approximately 700ng) and the quality of the gDNA. These limitations were examined by Illumina: they examined the effect of fragmentation by artificially digesting DNA and the effect of varying DNA input amount. They found that the overall call rate is relatively insensitive to quantity and fragmentation length across the entire range of DNA inputs (200ng down to 3ng). The total variation in the allelic ratios was minimally affected by input amount; however, the variation on the log R ration increased dramatically with a decrease in input amount and fragmented DNA. Variation of the log R ratio was greatly reduced when paired samples of similar quantity and fragment length were used. This indicates that small amount of samples are required for LOH or copy number analyses if performed in the paired analysis mode [71].

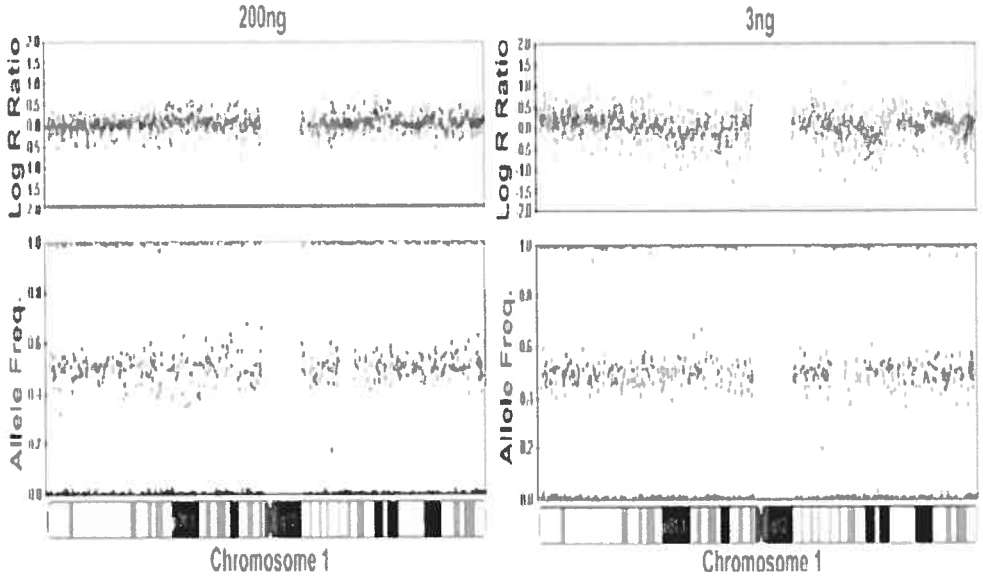


***Figure 51**: Effect of gDNA quantity on SNP-CGH data. Only limited amounts of sample are required for LOH or copy number analyses if performed in the paired analysis mode [71].*

Another possible limitation is the analysis of heterogeneous tumour samples since contamination of stromal components or lymphocytes can profound the distinction between true LOH and other mechanisms of allelic imbalance. High contamination levels were examined in the illumina report showing that 50% normal contamination, various aberrations can easily detected; however their interpretation remains ambiguous and at 75% normal contamination, aberrations are not visible [71]. This illustrates the possibility the study of tumors containing a rich stromal component. However, as we observed, detection of homozygous deletion is made challenging in the presence of contaminating normal gDNA since the allele frequency could be indistinguishable from normal. The resolution of the SNP array is not the cause of the inability of detecting the possible homozygous deletion which was shown by the FISH experiment. In fact, the resolution of the FISH probes used to examine the EBV-SMT was in the range of 550Kb which is comparable to the 109K Sentrix Human-1 SNP array higher resolution of 260Kb. This case illustrates that only highly pure gDNA is required to detect possible homozygous deletion.

4. SNP array can not detect translocations

A minor downfall of the SNP array is the inability to detect balanced translocations. Unbalanced translocations are detected by the SNP array as deletions or amplifications. As we have seen with the Infantile Myofibromatosis samples, there were no chromosomal abnormalities detected by the SNP array. However, if there was a possible balanced translocation event in these tumours, the SNP array is not capable of detecting it. The same possible limitation was seen in the EBV-SMT which also did not show any chromosomal aberrations. However, the easy, rapid, and robust SNP technology capable of identifying genome wide aberrations at ultrahigh resolution represents an important advance in clinical diagnostics. The minor down fall can be complimented with additional techniques, for example FISH, to give the best diagnosis to each patient suffering from rare paediatric tumours.

5. CGH will be replaced by SNP array

With Illumina's SNP-CGH, both copy number analysis obtained from intensity information and genotyping information by examining allelic imbalance can be obtained on the same platform. CGH does not provide any genotyping data. We presume that CGH arrays will eventually be replaced by SNP-CGH arrays due to their high performance level and their low cost.

6. Perspective, prediction and future studies

A. Importance of data base

Since each SNP can be precisely located in a deleted or amplified chromosomal region by Illumina Chromosome Browser, gene mapping would be greatly enhanced if the information obtained could be systematically incorporated into a database. Public data base would also facilitate research between laboratories and save precious time by sharing information.

B. Somatic cancer studies made easier

Many limitations have been overcome by the use of SNP array in the study of rare paediatric cancers. The simple necessity of tumoral DNA is sufficient for genotyping. No more cell culture will be necessary to achieve karyotypes for molecular diagnostics. Illumina's core BeadArray technology could be implemented in diagnostic laboratories and reduce the time of molecular tool optimization for rare paediatric tumors.

C. Prognostic and therapeutics ameliorated

In conclusion, we have shown that tumoral genotyping with SNP-CGH arrays offer several distinct advantages over conventional genetic tools in the study of paediatric cancers. By reducing the cost of large-scale genomics research, SNP arrays will enable the generation of information that will improve diagnostic and therapeutic approaches and, ultimately, enhance individual clinical outcomes.

VI. REFERENCES

1. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. **100**(1): p. 57-70.
2. Vogelstein, B. and K.W. Kinzler, *Cancer genes and the pathways they control.* Nat Med, 2004. **10**(8): p. 789-99.
3. Engle, L.J., C.L. Simpson, and J.E. Landers, *Using high-throughput SNP technologies to study cancer.* Oncogene, 2006. **25**(11): p. 1594-601.
4. White, K.A., et al., *Prohibitin mutations are uncommon in prostate cancer families linked to chromosome 17q.* Prostate Cancer Prostatic Dis, 2006. **9**(3): p. 298-302.
5. Kemp, Z.E., et al., *Evidence of linkage to chromosome 9q22.33 in colorectal cancer kindreds from the United Kingdom.* Cancer Res, 2006. **66**(10): p. 5003-6.
6. Zeng, Z., et al., *Family-based association analysis validates chromosome 3p21 as a putative nasopharyngeal carcinoma susceptibility locus.* Genet Med, 2006. **8**(3): p. 156-60.
7. Levitt, N.C. and I.D. Hickson, *Caretaker tumour suppressor genes that defend genome integrity.* Trends Mol Med, 2002. **8**(4): p. 179-86.
8. Knudson, A.G., Jr., *Mutation and cancer: statistical study of retinoblastoma.* Proc Natl Acad Sci U S A, 1971. **68**(4): p. 820-3.
9. Knudson, A.G., *Hereditary cancer: two hits revisited.* J Cancer Res Clin Oncol, 1996. **122**(3): p. 135-40.
10. Kinzler, K.W. and B. Vogelstein, *Cancer-susceptibility genes. Gatekeepers and caretakers.* Nature, 1997. **386**(6627): p. 761, 763.
11. Longo, L.D., *Classic pages in obstetrics and gynecology. The chromosome number in man. Joe Hin Tjio and Albert Levan. Hereditas, vol. 42, pp. 1-6, 1956.* Am J Obstet Gynecol, 1978. **130**(6): p. 722.
12. Porter, I.H., *The clinical side of cytogenetics.* J Reprod Med, 1976. **17**(1): p. 3-18.
13. Speicher, M.R. and N.P. Carter, *The new cytogenetics: blurring the boundaries with molecular biology.* Nat Rev Genet, 2005. **6**(10): p. 782-92.

14. Mitelman, F. and S. Heim, *Chromosome abnormalities in cancer.* Cancer Detect Prev, 1990. **14**(5): p. 527-37.

15. Aplan, P.D., *Causes of oncogenic chromosomal translocation.* Trends Genet, 2006. **22**(1): p. 46-55.

16. Franco, S., F.W. Alt, and J.P. Manis, *Pathways that suppress programmed DNA breaks from progressing to chromosomal breaks and translocations.* DNA Repair (Amst), 2006. **5**(9-10): p. 1030-41.

17. Albertson, D.G., et al., *Chromosome aberrations in solid tumors.* Nat Genet, 2003. **34**(4): p. 369-76.

18. Mitelman, F., B. Johansson, and F. Mertens, *Fusion genes and rearranged genes as a linear function of chromosome aberrations in cancer.* Nat Genet, 2004. **36**(4): p. 331-4.

19. Baak, J.P., et al., *Genomics and proteomics in cancer.* Eur J Cancer, 2003. **39**(9): p. 1199-215.

20. Eguchi, M., et al., *Fusion of ETV6 to neurotrophin-3 receptor TRKC in acute myeloid leukemia with t(12;15)(p13;q25).* Blood, 1999. **93**(4): p. 1355-63.

21. Mertens, F., et al., *Chromosomal imbalance maps of malignant solid tumors: a cytogenetic survey of 3185 neoplasms.* Cancer Res, 1997. **57**(13): p. 2765-80.

22. Marculescu, R., et al., *Recombinase, chromosomal translocations and lymphoid neoplasia: Targeting mistakes and repair failures.* DNA Repair (Amst), 2006. **5**(9-10): p. 1246-58.

23. Weinberg, R.A., *Tumor suppressor genes.* Science, 1991. **254**(5035): p. 1138-46.

24. Savelyeva, L. and M. Schwab, *Amplification of oncogenes revisited: from expression profiling to clinical application.* Cancer Lett, 2001. **167**(2): p. 115-23.

25. Iehara, T., et al., *MYCN gene amplification is a powerful prognostic factor even in infantile neuroblastoma detected by mass screening.* Br J Cancer, 2006. **94**(10): p. 1510-5.

26. Zawrocki, A. and W. Biernat, *Epidermal growth factor receptor in glioblastoma.* Folia Neuropathol, 2005. **43**(3): p. 123-32.

27. Sait, S.N., et al., *Double minute chromosomes in acute myeloid leukemia and myelodysplastic syndrome: identification of new amplification regions by fluorescence in situ hybridization and spectral karyotyping.* Genes Chromosomes Cancer, 2002. **34**(1): p. 42-7.

28. Surace, C., et al., *Fluorescent in situ hybridization (FISH) reveals frequent and recurrent numerical and structural abnormalities in hepatoblastoma with no informative karyotype.* Med Pediatr Oncol, 2002. **39**(5): p. 536-9.

29. Garnis, C., T.P. Buys, and W.L. Lam, *Genetic alteration and gene expression modulation during cancer progression.* Mol Cancer, 2004. **3**: p. 9.

30. Liu, J.Y., et al., *Detection of human chromosomal abnormalities using a new technique combining 4',6-diamidino-2-phenyl-indole staining and image analysis.* Clin Genet, 2006. **69**(1): p. 65-71.

31. Bezrookove, V., et al., *Individuals with abnormal phenotype and normal G-banding karyotype: improvement and limitations in the diagnosis by the use of 24-colour FISH.* Hum Genet, 2000. **106**(4): p. 392-8.

32. Lillington, D.M., et al., *Detection of chromosome abnormalities pre-high-dose treatment in patients developing therapy-related myelodysplasia and secondary acute myelogenous leukemia after treatment for non-Hodgkin's lymphoma.* J Clin Oncol, 2001. **19**(9): p. 2472-81.

33. Osoegawa, K., et al., *A bacterial artificial chromosome library for sequencing the complete human genome.* Genome Res, 2001. **11**(3): p. 483-96.

34. Osoegawa, K., et al., *An improved approach for construction of bacterial artificial chromosome libraries.* Genomics, 1998. **52**(1): p. 1-8.

35. Gershon, D., *DNA microarrays: more than gene expression.* Nature, 2005. **437**(7062): p. 1195-8.

36. Yuan, E., et al., *Genomic profiling maps loss of heterozygosity and defines the timing and stage dependence of epigenetic and genetic events in Wilms' tumors.* Mol Cancer Res, 2005. **3**(9): p. 493-502.

37. Dallapiccola, B., *Cytogenetics of Mendelian mutations associated with cancer proneness.* Cancer Genet Cytogenet, 1987. **26**(1): p. 85-94.

38. Vrieling, H., *Mitotic maneuvers in the light.* Nat Genet, 2001. **28**(2): p. 101-2.

39. Zhao, X., et al., *An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays.* Cancer Res, 2004. **64**(9): p. 3060-71.

40. Tomlinson, I.P., M.B. Lambros, and R.R. Roylance, *Loss of heterozygosity analysis: practically and conceptually flawed?* Genes Chromosomes Cancer, 2002. **34**(4): p. 349-53.

41. Perri, P., et al., *Restriction fragment length polymorphism analysis reveals different allele frequency and a linkage disequilibrium at locus D1S94 in neuroblastoma patients.* Eur J Cancer, 1997. **33**(12): p. 1949-52.

42. el-Naggar, A.K., et al., *Polymerase chain reaction-based restriction fragment length polymorphism analysis of the short arm of chromosome 3 in primary head and neck squamous carcinoma.* Cancer, 1993. **72**(3): p. 881-6.

43. Meltzer, S.J., et al., *Reduction to homozygosity involving p53 in esophageal cancers demonstrated by the polymerase chain reaction.* Proc Natl Acad Sci U S A, 1991. **88**(11): p. 4976-80.

44. Ellegren, H., *Microsatellites: simple sequences with complex evolution.* Nat Rev Genet, 2004. **5**(6): p. 435-45.

45. Rigaud, G., et al., *Allelotype of pancreatic acinar cell carcinoma.* Int J Cancer, 2000. **88**(5): p. 772-7.

46. Hampton, G.M., et al., *Simultaneous assessment of loss of heterozygosity at multiple microsatellite loci using semi-automated fluorescence-based detection: subregional mapping of chromosome 4 in cervical carcinoma.* Proc Natl Acad Sci U S A, 1996. **93**(13): p. 6704-9.

47. Wang, Z.C., et al., *Genome-Wide Analysis for Loss of Heterozygosity in Primary and Recurrent Phyllodes Tumor and Fibroadenoma of Breast using Single Nucleotide Polymorphism Arrays.* Breast Cancer Res Treat, 2006. **97**(3): p. 301-9.

48. Nishimura, T., et al., *Genotype stability and clonal evolution of hepatocellular carcinoma assessed by autopsy-based genome-wide microsatellite analysis.* Cancer Genet Cytogenet, 2005. **161**(2): p. 164-9.

49. Wong, K.K., et al., *Allelic imbalance analysis by high-density single-nucleotide polymorphic allele (SNP) array with whole genome amplified DNA.* Nucleic Acids Res, 2004. **32**(9): p. e69.

50. Zheng, H.T., et al., *Loss of heterozygosity analyzed by single nucleotide polymorphism array in cancer.* World J Gastroenterol, 2005. **11**(43): p. 6740-4.

51. Janne, P.A., et al., *High-resolution single-nucleotide polymorphism array and clustering analysis of loss of heterozygosity in human lung cancer cell lines.* Oncogene, 2004. **23**(15): p. 2716-26.

52. Hoque, M.O., et al., *Genome-wide genetic characterization of bladder cancer: a comparison of high-density single-nucleotide polymorphism arrays and PCR-based microsatellite analysis.* Cancer Res, 2003. **63**(9): p. 2216-22.

53. Lindblad-Toh, K., et al., *Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays.* Nat Biotechnol, 2000. **18**(9): p. 1001-5.

54. Andersen, C.L., et al., *Frequent occurrence of uniparental disomy in colorectal cancer.* Carcinogenesis, 2006.

55. Walker, B.A., et al., *Integration of global SNP-based mapping and expression arrays reveals key regions, mechanisms, and genes important in the pathogenesis of multiple myeloma.* Blood, 2006. **108**(5): p. 1733-43.

56. Maris, J.M., et al., *Region-specific detection of neuroblastoma loss of heterozygosity at multiple loci simultaneously using a SNP-based tag-array platform.* Genome Res, 2005. **15**(8): p. 1168-76.

57. Huang, J., et al., *Whole genome DNA copy number changes identified by high density oligonucleotide arrays.* Hum Genomics, 2004. **1**(4): p. 287-99.

58. Zhou, X., et al., *Concurrent analysis of loss of heterozygosity (LOH) and copy number abnormality (CNA) for oral premalignancy progression using the Affymetrix 10K SNP mapping array.* Hum Genet, 2004. **115**(4): p. 327-30.

59. Bignell, G.R., et al., *High-resolution analysis of DNA copy number using oligonucleotide microarrays.* Genome Res, 2004. **14**(2): p. 287-95.

60. Wang, Z.C., et al., *Loss of heterozygosity and its correlation with expression profiles in subclasses of invasive breast cancers.* Cancer Res, 2004. **64**(1): p. 64-71.

61. Dumur, C.I., et al., *Genome-wide detection of LOH in prostate cancer using human SNP microarray technology.* Genomics, 2003. **81**(3): p. 260-9.

62. Paige, A.J., *Redefining tumour suppressor genes: exceptions to the two-hit hypothesis.* Cell Mol Life Sci, 2003. **60**(10): p. 2147-63.

63. Oligny, L.L., *Cancer and epigenesis: a developmental perspective.* Adv Pediatr, 2003. **50**: p. 59-80.

64. Feinberg, A.P., R. Ohlsson, and S. Henikoff, *The epigenetic progenitor origin of human cancer.* Nat Rev Genet, 2006. **7**(1): p. 21-33.

65. Schulz, R., et al., *Chromosome-wide identification of novel imprinted genes using microarrays and uniparental disomies.* Nucleic Acids Res, 2006. **34**(12): p. e88.

66. Gyapay, G., et al., *The 1993-94 Genethon human genetic linkage map.* Nat Genet, 1994. **7**(2 Spec No): p. 246-339.

67. Ting, J.C., et al., *Analysis and visualization of chromosomal abnormalities in SNP data with SNPscan.* BMC Bioinformatics, 2006. **7**(1): p. 25.

68. Teh, M.T., et al., *Genomewide single nucleotide polymorphism microarray mapping in basal cell carcinomas unveils uniparental disomy as a key somatic event.* Cancer Res, 2005. **65**(19): p. 8597-603.

69. Thompson, E.R., et al., *Whole genome SNP arrays using DNA derived from formalin-fixed, paraffin-embedded ovarian tumor tissue.* Hum Mutat, 2005. **26**(4): p. 384-9.

70. Steemers, F.J. and K.L. Gunderson, *Illumina, Inc.* Pharmacogenomics, 2005. **6**(7): p. 777-82.

71. Peiffer, D.A., et al., *High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping.* Genome Res, 2006.

72. Gunderson, K.L., et al., *A genome-wide scalable SNP genotyping assay using microarray technology.* Nat Genet, 2005. **37**(5): p. 549-54.

73. Yeh, A., et al., *Chromosome arm 16q in Wilms tumors: unbalanced chromosomal translocations, loss of heterozygosity, and assessment of the CTCF gene.* Genes Chromosomes Cancer, 2002. **35**(2): p. 156-63.

74. Mummert, S.K., V.A. Lobanenkov, and A.P. Feinberg, *Association of chromosome arm 16q loss with loss of imprinting of insulin-like growth*

factor-II in Wilms tumor. Genes Chromosomes Cancer, 2005. **43**(2): p. 155-61.

75. Strefford, J.C., et al., *A combination of molecular cytogenetic analyses reveals complex genetic alterations in conventional renal cell carcinoma.* Cancer Genet Cytogenet, 2005. **159**(1): p. 1-9.
76. Marsit, C.J., et al., *Loss of heterozygosity of chromosome 3p21 is associated with mutant TP53 and better patient survival in non-small-cell lung cancer.* Cancer Res, 2004. **64**(23): p. 8702-7.
77. Yoon, J.H., R. Dammann, and G.P. Pfeifer, *Hypermethylation of the CpG island of the RASSF1A gene in ovarian and renal cell carcinomas.* Int J Cancer, 2001. **94**(2): p. 212-7.
78. Chan, M.W., et al., *Frequent hypermethylation of promoter region of RASSF1A in tumor tissues and voided urine of urinary bladder cancer patients.* Int J Cancer, 2003. **104**(5): p. 611-6.
79. Peng, H., T. Zhao, and K.T. Yao, *Expression of RASSF1A gene in nasopharyngeal carcinoma.* Di Yi Jun Yi Da Xue Xue Bao, 2003. **23**(7): p. 673-6.
80. Perotti, D., et al., *Germline mutations of the POU6F2 gene in Wilms tumors with loss of heterozygosity on chromosome 7p14.* Hum Mutat, 2004. **24**(5): p. 400-7.
81. Sossey-Alaoui, K., et al., *Molecular characterization of a 7p15-21 homozygous deletion in a Wilms tumor.* Genes Chromosomes Cancer, 2003. **36**(1): p. 1-6.
82. Perotti, D., et al., *Refinement within single yeast artificial chromosome clones of a minimal region commonly deleted on the short arm of chromosome 7 in Wilms tumours.* Genes Chromosomes Cancer, 2001. **31**(1): p. 42-7.
83. Powlesland, R.M., et al., *Loss of heterozygosity at 7p in Wilms' tumour development.* Br J Cancer, 2000. **82**(2): p. 323-9.
84. Grundy, R.G., et al., *Loss of heterozygosity for the short arm of chromosome 7 in sporadic Wilms tumour.* Oncogene, 1998. **17**(3): p. 395-400.

85. Satoh, Y., et al., *Genetic and epigenetic alterations on the short arm of chromosome 11 are involved in a majority of sporadic Wilms' tumours.* Br J Cancer, 2006. **95**(4): p. 541-547.

86. Lo Muzio, L., et al., *Myofibroma/myofibromatosis is the most common fibrous proliferation of infancy and childhood.* Oral Oncol, 2000. **36**(1): p. 144.

87. Hatzidaki, E., et al., *Infantile myofibromatosis with visceral involvement and complete spontaneous regression.* J Dermatol, 2001. **28**(7): p. 379-82.

88. Stenzel, P. and S. Fitterer, *Gastrointestinal multicentric infantile myofibromatosis: characteristic histology on rectal biopsy.* Am J Gastroenterol, 1989. **84**(9): p. 1115-9.

89. Coffin, C.M., et al., *Congenital generalized myofibromatosis: a disseminated angiocentric myofibromatosis.* Pediatr Pathol Lab Med, 1995. **15**(4): p. 571-87.

90. Molnar, P., et al., *Aggressive infantile myofibromatosis: report of a case of a clinically progressive congenital multiple fibromatosis.* Med Pediatr Oncol, 1986. **14**(6): p. 332-7.

91. Sirvent, N., et al., *Monosomy 9q and trisomy 16q in a case of congenital solitary infantile myofibromatosis.* Virchows Arch, 2004. **445**(5): p. 537-40.

92. Ikediobi, N.I., et al., *Infantile myofibromatosis: support for autosomal dominant inheritance.* J Am Acad Dermatol, 2003. **49**(2 Suppl Case Reports): p. S148-50.

93. Zand, D.J., et al., *Autosomal dominant inheritance of infantile myofibromatosis.* Am J Med Genet, 2004. **126A**(3): p. 261-6.

94. Jennings, T.A., et al., *Infantile myofibromatosis. Evidence for an autosomal-dominant disorder.* Am J Surg Pathol, 1984. **8**(7): p. 529-38.

95. Bracko, M., L. Cindro, and R. Golouh, *Familial occurrence of infantile myofibromatosis.* Cancer, 1992. **69**(5): p. 1294-9.

96. Venencie, P.Y., et al., *Infantile myofibromatosis. Report of two cases in one family.* Br J Dermatol, 1987. **117**(2): p. 255-9.

97. Narchi, H., *Four half-siblings with infantile myofibromatosis: a case for autosomal-recessive inheritance.* Clin Genet, 2001. **59**(2): p. 134-5.

98. Stenman, G., et al., *del(6)(q12q15) as the sole cytogenetic anomaly in a case of solitary infantile myofibromatosis.* Oncol Rep, 1999. **6**(5): p. 1101-4.

99. Cheung, V.G., et al., *Integration of cytogenetic landmarks into the draft sequence of the human genome.* Nature, 2001. **409**(6822): p. 953-8.

100. Paternoster, S.F., et al., *A new method to extract nuclei from paraffin-embedded tissue to study lymphomas using interphase fluorescence in situ hybridization.* Am J Pathol, 2002. **160**(6): p. 1967-72.

101. Boudjemaa, S., et al., *Brain involvement in multicentric Epstein-Barr virus-associated smooth muscle tumours in a child after kidney transplantation*, in *Virchows Arch.* 2004. p. 387-91.

102. Debiec-Rychter, M., et al., *Complex genomic rearrangement of ALK loci associated with integrated human Epstein-Barr virus in a post-transplant myogenic liver tumor.* Am J Pathol, 2003. **163**(3): p. 913-22.

103. Willman, C.L., et al., *Langerhans'-cell histiocytosis (histiocytosis X)--a clonal proliferative disease.* N Engl J Med, 1994. **331**(3): p. 154-60.

104. Shih, L.Y., et al., *Clonality analysis using X-chromosome inactivation patterns by HUMARA-PCR assay in female controls and patients with idiopathic thrombocytosis in Taiwan.* Exp Hematol, 2001. **29**(2): p. 202-8.

105. Allen, R.C., et al., *Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation*, in *Am J Hum Genet.* 1992. p. 1229-39.