

Analyse des résultats de puce à ADN

1 OBJECTIF

Ce document décrit les critères qualité à prendre en compte avant toute analyse d'un profil génomique ainsi que les règles de base d'analyse pour les puces à ADN Agilent et Affymetrix (Oncoscan CNV et Cytoscan HD).

2 DOMAINE D'APPLICATION

Puce Agilent SurePrint G3 Human CGH 8x60k uniquement :

- 👉 Analyse de l'index génomique dans les GIST, protocole GIGIST: puce Agilent
- 👉 Analyse de l'index génomique dans les LMS/LM/STUMP de l'utérus
- 👉 Analyse de l'index génomique dans les synoviosarcomes
- 👉 Analyse de l'index génomique d'une tumeur

Puce Affymetrix uniquement :

- 👉 Protocole Safir 02 : puces Affymetrix
- 👉 Tumeurs rhabdoïdes

Puce Agilent ou Affymetrix selon le cas :

- 👉 Protocole Inpakt
- 👉 Analyse du profil de deux tumeurs chez un même patient (tumeurs métachrones)
- 👉 Tumeur bénigne ou maligne
- 👉 Lipome à cellules fusiformes
- 👉 Vérification d'un résultat de NGS (présence d'une amplification ou délétion suspectée)
- 👉 Cas équivoques en FISH pour la recherche d'amplification des gènes *MDM2* ou *HER2*
- 👉 Recherche de point de cassure dans des gènes décrits comme pouvant être impliqués dans des translocations
- 👉 Recherche d'altérations de liste de gènes définis

3 PERSONNEL CONCERNÉ

- 👉 Techniciens et ingénieurs de Pathologie Moléculaire
- 👉 Médecins et internes de Pathologie Moléculaire

4 DOCUMENTS ASSOCIES

- 👉 PAM-PROC-0020 Réalisation d'un examen de CGH ou puce à ADN
- 👉 PAM-FITC-0054 Critères d'acceptation de prélèvements pour une analyse sur puce à ADN
- 👉 PAM-PROC-0016 Gestion du travail technique dans le SIL Ariane en Pathologie Moléculaire
- 👉 PAM-PROC-0011 Gestion des contaminations en PAM
- 👉 PAM-MOP-0076 Protocole de CGH Agilent puce SurePrint G3 Human CGH 8x60k
- 👉 PAM-MOP-0073 Protocole puces à ADN Cytoscan HD
- 👉 PAM-MOP-0074 Protocole puces à ADN Oncoscan CNV
- 👉 PAM-MOP-0072 Utilisation du logiciel Agilent Cytogenomics
- 👉 PAM-MOP-0075 Utilisation du logiciel Affymetrix Chromosome analysis suite ChAS
- 👉 PAM-FQ-0170 Feuille de travail PCR SRY
- 👉 PAM-DEXT-0130 Manuel fournisseur Affymetrix Logiciel ChAS
- 👉 PAM-DEXT-0129 Manuel fournisseur Oncoscan console
- 👉 PAM-DEXT-0124 Manuel fournisseur Agilent Logiciel Cytogenomics
- 👉 PAM-DEXT-0131 Document fournisseur exemples d'analyse de résultats Oncoscan
- 👉 PAM-DEXT-0132 Document fournisseur Affymetrix analyse de la différence allélique
- 👉 PAM-DEXT-0133 Document fournisseur exemples d'analyse de résultats de puce Cytoscan HD
- 👉 PAM-MOP-0044 Synthèse analytique de résultats de puce à ADN dans le SIL ariane
- 👉 PAM-PROC-0017 Création et diffusion des comptes-rendus de Pathologie Moléculaire dans le SIL Ariane
- 👉 BPAT-PROC-0030 Non-conformités et incidents dans Ariane SX

5 MATERIEL UTILISE

Logiciel Agilent : CytoGenomics

Logiciel Affymetrix : Chromosome Analysis Suite (ChAS)

6 DEROULEMENT DES ACTIVITES

6.1 Puce Agilent SurePrint G3 Human CGH 8x60k

6.1.1 Description de la puce 8X60k

La technique utilisée sur ces puces est une hybridation génomique comparative (CGH). L'ADN d'un échantillon tumoral est marqué avec un fluorochrome et est co-hybridé sur une puce à ADN avec de l'ADN normal marqué avec un autre fluorochrome. La puce est ensuite scannée et la fluorescence des 2 fluorochromes au niveau de chaque sonde de la puce est mesurée. Les données obtenues sont les ratios de fluorescence ADN tumoral/ADN normal en log 2 pour chaque sonde de la puce.

Cette puce présente environ 60 000 sondes et ces dernières correspondent à des oligonucléotides de 60 pb.

La puce permet d'hybrider sur une même lame jusqu'à 8 échantillons et présente une sonde toutes les 41 kb de manière générale et toutes les 33 kb dans les gènes Refseq.

6.1.2 Vérifications préalables

- Se référer au document **PAM-MOP-0072** pour l'utilisation du logiciel
- Vérifier la conformité des témoins techniques : si un témoin technique (blanc d'extraction et/ou CIQ) est non conforme se référer à la procédure **PAM-PROC-0011**
- Le **technicien** à la fin du scan doit réaliser ces vérifications et le **médecin** qui analyse les profils doit également les refaire avant toute analyse : aucun cas ne doit être analysé si ces deux vérifications indépendantes n'ont pas été réalisées

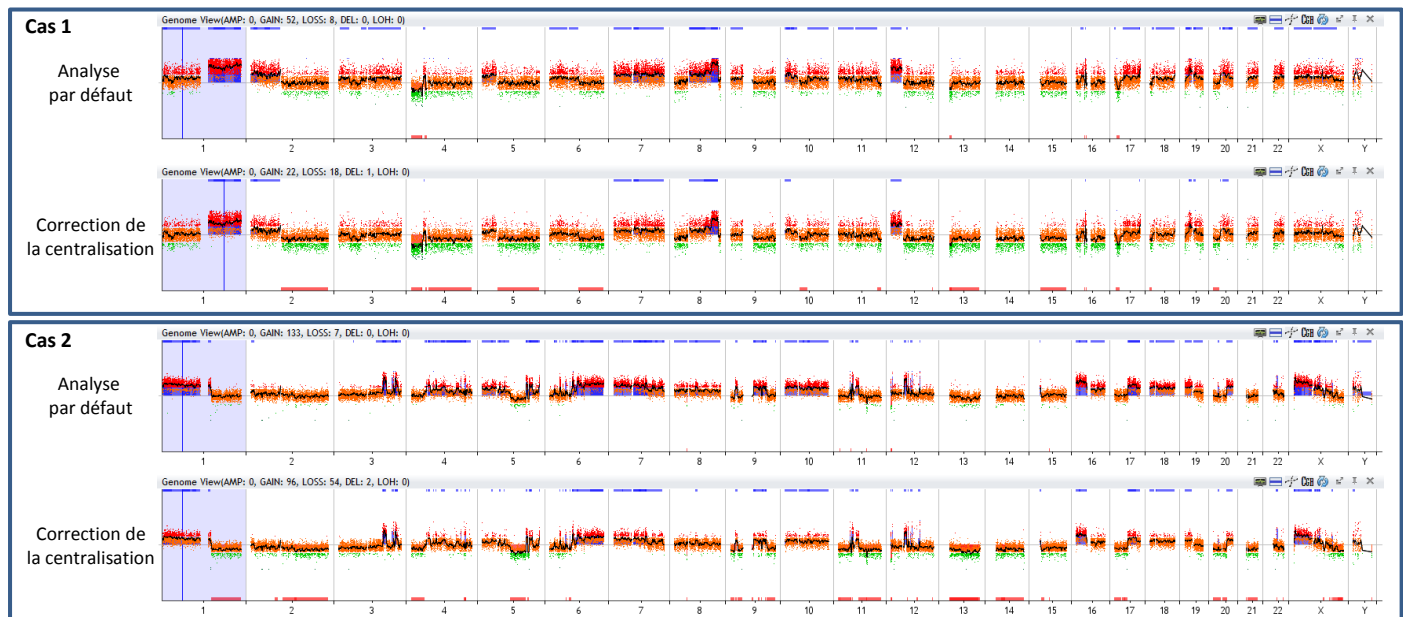
6.1.2.1 Qualité de la technique

- Ne pas interpréter une puce si le **DLRS** est supérieur à **0,4** en FFPE et **0,3** en congélation

6.1.2.2 Centralisation du cas

- Parfois l'analyse par défaut n'est pas la plus adaptée si la tumeur présente beaucoup d'altérations génomiques
- Si la tumeur présente une grande majorité d'altérations dans le même sens (pratiquement que des gains/amplifications ou que des pertes) il est possible que la centralisation du cas ne soit pas optimale
- Pour réanalyser le cas avec une correction de la centralisation se référer au document : **PAM-MOP-0072**

Voici deux exemples de cas pour lesquels une correction de la centralisation a été apportée :



6.1.2.3 Vérification du sexe obtenu en CGH

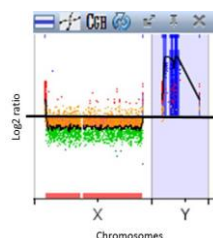
- Vérifier si le **sexe** en CGH correspond bien au sexe du patient, regarder sur le profil génomique les statuts des chromosomes X et Y :

- Si l'ADN de référence est celui d'un homme :

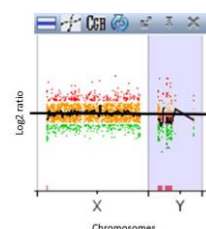
- ✓ La tumeur doit présenter si c'est celle d'un **homme** : un statut normal pour les chromosomes X et Y
- ✓ La tumeur doit présenter si c'est celle d'une **femme** : un statut gagné pour le chromosome X et une perte pour le chromosome Y

- Si l'ADN de référence est celui d'une femme :

- ✓ La tumeur doit présenter si c'est celle d'un **homme** : une perte d'un chromosome X et un gain/amplification pour le chromosome Y :



- ✓ La tumeur doit présenter si c'est celle d'une **femme** : un statut normal pour les chromosomes X et Y :



Il y a visualisation d'un statut normal pour le chromosome Y chez les femmes car ce sont les régions pseudo-autosomales de ce chromosome qui sont ciblées sur la puce.

Attention rester conscient que la tumeur d'une femme peut perdre un chromosome X tout comme la tumeur d'un homme peut perdre un chromosome Y ou gagner un ou plusieurs chromosomes X. Toujours vérifier tous les sexes de tous les échantillons à la suite et interpréter la série dans sa globalité.

S'il y a un problème de cohérence entre le sexe obtenu en CGH et le sexe du patient, en référer à un supérieur et :

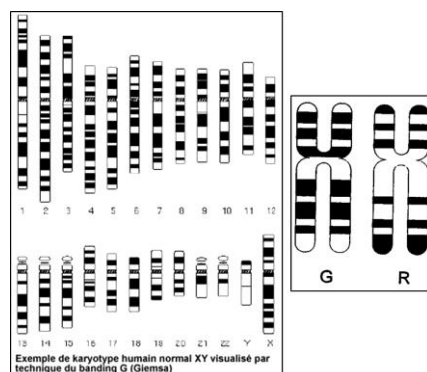
- vérifier s'il n'y a pas eu d'erreur de saisi du sexe dans le SIL à réception des prélèvements
 - vérifier l'identité des tubes d'ADN initiaux et leur ordre d'utilisation
 - vérifier l'ordre des tubes conservés tout au long de la technique
 - faire une PCR SRY sur les tubes d'ADN initiaux pour vérifier le sexe de l'échantillon extrait afin de voir si le problème pourrait venir de la technique de CGH ou s'il est antérieur (**PAM-FQ-0170**)
- Tracer toutes les vérifications faites dans les feuilles de série (**PAM-FQ-0171 à PAM-FQ-0174**) et indiquer l'origine du problème si elle a été identifiée
 - Faire une non-conformité ou une déclaration d'incident dans le SIL selon le problème, avertir immédiatement le pathologiste demandeur et voir la marche à suivre avec l'ingénieur, un médecin de PAM et/ou le pathologiste demandeur selon l'erreur (**BPAT-PROC-0030**)

6.1.2.4 Effet bandes R

Les **bandes R** sont des bandes chromosomiques mises en évidence par marquage/coloration des chromosomes métaphasiques sur des étalements. Elles sont notamment obtenues par dénaturation thermique de l'ADN puis coloration au Giemsa (RHG). Elles sont riches en bases GC et en gènes.

Les **bandes G** quant à elles sont couramment obtenues par digestion enzymatique des préparations chromosomiques par la trypsine puis coloration au Giemsa (GTG). Elles sont riches en bases AT et sont pauvres en gènes.

Le marquage au Giemsa permet de révéler des bandes noires marquées et des bandes plus claires moins marquées qui diffèrent selon le prétraitement des chromosomes. La disposition des bandes R est l'inverse de celle qui est obtenue en bandes G (bandes R pour Reverse) :



6.1.3 Analyse

- L'analyse ne doit pas être réalisée si les vérifications préalables n'ont pas été faites par le technicien et le médecin de manière indépendante et s'il y a le moindre doute quant à un des points mentionnés dans la partie au-dessus
- Analyse : **toujours interpréter dans le contexte d'une région donnée, une seule sonde altérée peut-être artefactuelle**
 - ✓ Un statut **normal** correspond à un ratio en log2 compris entre -0,5 et 0,5
 - ✓ Un **gain** correspond à un ratio en log2 compris entre 0,6 et 1,9
 - ✓ Une **amplification** correspond à un ratio en log2 supérieur à 2
 - ✓ Une **perte hétérozygote** correspond à un ratio en log2 compris entre -0,6 et -1,5
 - ✓ Une **perte homozygote** est suspectée quand le ratio en log2 est inférieur à -1,6
- Attention pour des altérations de quelques sondes sur une région il faut vérifier qu'il ne s'agisse pas d'une région de type CNV (copy number variation) qui sont des polymorphismes de nombre de copies et qui donc peuvent différer entre le patient et l'ADN de référence utilisé

6.1.4 Calcul d'index génomique

- **Calculer le GI = (nb altérations)²/(nb de chromosomes impliqués)**

Compter comme altérations uniquement les altérations détectées par le logiciel (identifiées par une barre rouge ou bleue) à l'exception de :

- ✓ Ne pas compter les altérations sur les chromosomes sexuels
- ✓ Attention aux CNV polymorphiques (ex gènes ADAM) : à ne pas prendre en compte
- ✓ Attention aux altérations télomériques ou péricentromériques pouvant être artefactuelles
- ✓ Attention : si le centromère est inclus dans la perte on doit compter 2 altérations comme le logiciel
- Nb de sondes altérées pour définir une altération :
 - **Puce 8X60k** : au moins 100 sondes altérées
- Seuil de GI selon les tumeurs:
 - ❖ **GIST** : **GI < 10** : tumeur de bon pronostic - **GI ≥ 10** : tumeur de plus mauvais pronostic
 - ❖ **Tumeurs utérines** : **GI < 10** : tumeur de bon pronostic - **GI ≥ 10** : tumeur de plus mauvais pronostic
 - ❖ **Synoviosarcome** : **GI < 1** : tumeur de bon pronostic - **GI ≥ 1** : tumeur de plus mauvais pronostic
 - ❖ **Pour tout autre type de tumeur** : aucun seuil n'a encore été publié, le GI n'est donné qu'à titre indicatif

NB : altérations fréquentes dans les GIST : perte des chromosomes 14 – 22 et 1p

6.2 Puces Affymetrix Cytoscan HD et Oncoscan CNV

6.2.1 Description des puces

6.2.1.1 Cytoscan HD

Elle contient au total environ **6.5 million** de sondes: 1 sonde par allèle en triplicat pour 750 000 sondes SNPs et 1,9 million de sondes non-polymorphes. Les sondes sont des oligonucléotides de **25 pb** synthétisées sur la lame par photolithographie, chaque sonde correspond à des milliers d'oligonucléotides identiques. La puce est en silicium. La puce couvre à la fois les gènes constitutionnels et les gènes décrits comme étant impliqués dans des cancers avec :

- Une couverture intragénique globale de 1 marqueur/880 bases
- ClinGen (ISCA) couverture constitutionnelle de 1 marqueur/384 bases
- Une couverture complète des gènes impliqués dans des cancers de 1 marqueur/553 bases
- 12,000 gènes OMIM® avec 1 marqueur/659 bases
- >36,000 RefSeq gènes avec 1 marqueur/880 bases
- Une couverture non génique de 1 marqueur/1337 bases à travers le génome pour les points de cassure

Une couverture globale (gène and non-gène backbone) de 1 marqueur/1,148 bases

6.2.1.2 Oncoscan CNV

Elle contient au total **217454** sondes pratiquement toutes sont polymorphes (moins de 1000 sondes non polymorphes). Les sondes sont des oligonucléotides de **25 pb** synthétisées sur la lame par photolithographie, chaque sonde correspond à des milliers d'oligonucléotides identiques. La puce est en silicium.

Elle présente une résolution de 50 kb à 125 kb sur 900 gènes décrits comme étant impliqués dans des cancers.

En dehors de ces gènes:

- 90% du génome présente une résolution de 300-310 kb
- 97% du génome présente une résolution d'au moins 380 kb

La résolution des LOH à travers le génome est ≤ 10 MB. Si la fraction de cellules présentant l'anomalie est grande la résolution augmente jusqu'à 3 à 5 MB.

6.2.2 Principe de l'analyse

La technique utilisée avec les puces Oncoscan et Cytoscan n'est pas une hybridation génomique comparative, à la différence de la CGH Agilent, car seul est hybridé l'ADN tumoral sur la puce.

Les données obtenues sur la puce sont comparées via des algorithmes du logiciel ChAS à des bibliothèques de référence obtenues sur des tissus sains :

➤ **Pour la puce Cytoscan HD :**

Les fichiers du modèle de référence contiennent les données obtenues sur 380 puces qui ont été réalisées par 9 opérateurs différents par lots d'environ 48 échantillons uniques passés chacun deux fois avec une randomisation du placement des échantillons dans les plaques de PCR et une randomisation des réactifs et équipements utilisés.

Les 380 échantillons consistent en:

- ✓ 284 échantillons HapMap, incluant au moins un réplicat de chacun des 270 échantillons HapMap: 90 échantillons de chacun des groupes ethniques suivants Yoruban, Asiatique et Caucasien dérivés d'ADN de lignées cellulaires provenant du Coriell Institute of Medical Research
- ✓ 96 échantillons d'ADN extrait du sang d'hommes et de femmes phénotypiquement sains (BioServe Biotechnologies)

HapMap: le projet HapMap est un projet international (Canada, USA, Japon, Chine, Nigéria, UK) de génotypage de SNP visant à définir une carte d'haplotypes de différents groupes ethniques.

➤ **Pour la puce Oncoscan CNV :**

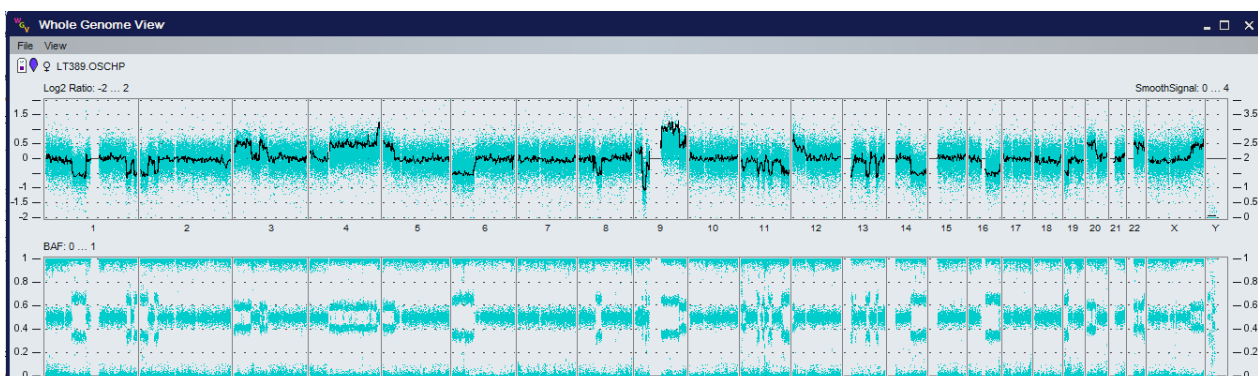
Les fichiers du modèle de référence contiennent les données obtenues sur environ 400 tissus FFPE normaux et tissus FFPE normaux adjacents à des tumeurs, provenant d'une 20^{aine} de sources couvrant une large sélection de zones géographiques, de sites collecteurs, d'âge de blocs, de types de cancer et de sexe.

Ainsi le **logiciel ChAS** après analyse des données obtenues sur la puce (Cytoscan ou Oncoscan) comparativement à la librairie de référence, permet de visualiser et de résumer les aberrations chromosomiques présentes tout au long du génome d'une tumeur. Parmi les aberrations visualisables on trouve les gains ou pertes de chromosomes ou de parties de chromosomes, la ploïdie, le mosaïcisme, les pertes d'hétérozygotie.

En effet le logiciel permet à la fois de voir :

En vue globale sur le génome :

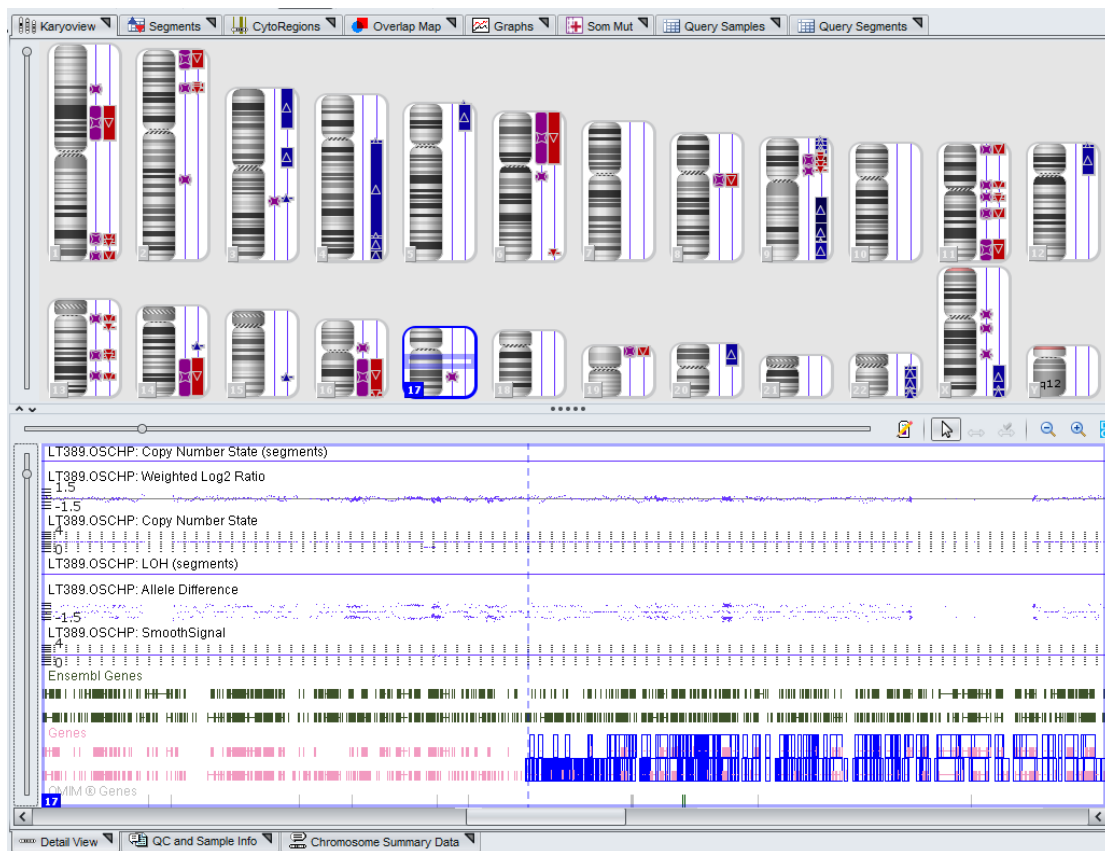
- Le Ratio en log2
- La fréquence allélique de l'allèle B (BAF : B allele frequency) pour **Oncoscan** et l'allele difference pour **Cytoscan**



Il est possible d'obtenir une vue des chromosomes avec 3 couleurs différentes : cliquer sur « View » dans la fenêtre de Whole Genome view et sélectionner « Colors », puis sélectionner « Alternating colors » et cliquer sur OK

En vue par chromosome :

- Le ratio en log2
- Le ratio en log2 lissé
- La fréquence allélique de l'allèle B (BAF : B allele frequency)
- Le nombre de copies pour chaque région de la puce (copy number)
- La perte d'hétérozygotie par région (LOH)
- La différence allélique
- Les gènes présents dans chaque région



6.2.3 Vérifications préalables

- Se référer au document **PAM-MOP-0075** pour l'utilisation du logiciel
- Vérifier la conformité des témoins techniques : si un témoin technique (blanc d'extraction, témoin négatif et/ou témoin positif) est non conforme se référer à la procédure **PAM-PROC-0011**
- Le **technicien** à la fin du scan doit réaliser ces vérifications et le **médecin** qui analyse les profils doit également les refaire avant toute analyse : aucun cas ne doit être analysé si ces deux vérifications indépendantes n'ont été réalisées

6.2.3.1 Critères qualité

6.2.3.1.1 Critères généraux

❖ Description des critères qualité:

1/ **MAPD** : correspond à la mesure globale de la variation de toutes les sondes de la puce sur l'ensemble du génome. Elle représente la médiane de distribution des différences entre les ratio en log2 de sondes adjacentes. Elle reflète la dispersion des points autour de la ligne de base du ratio en log2.

Cette variabilité peut avoir plusieurs sources:

- La variabilité intrinsèque au matériel de départ, la préparation du cocktail d'hybridation, la puce ou bien le scanner
- Une variabilité apparente due au fait que la référence peut présenter des différences systématiques par rapport à l'échantillon sur cette puce

Quelle que soit la source de cette variabilité elle diminue la qualité du compte de nombre de copies et peut conduire à des fausses altérations positives ou négatives

2/ **SNPQC**: correspond à une estimation de la distribution des allèles homozygotes AA, hétérozygotes AB et homozygotes BB et elle calcule une distance entre ces génotypes. Plus la séparation de ces distributions est grande plus il sera facile de définir le génotype basé sur sa position. Un faible SNPQC indique que la qualité des données de SNP est compromise à cause d'un fort bruit de fond sur la puce qui compromet sa qualité globale et la fiabilité de l'interprétation. Les prélèvements tumoraux peuvent avoir des SNPQC faibles du fait de la présence de nombreux remaniements, de leur hétérogénéité, de la contamination en cellules normales ou la présence de nombreux sous-clones.

ndSNPQC : SNP quality control calculer dans le cadre de l'analyse **Normal diploïd ou de prélèvement FFPE en Oncoscan**: son calcul n'est fait que sur les sondes qui apparaissent normales dans les régions diploïdes dans l'échantillon

3/ **Waviness SD**: correspond à une mesure globale de la variation des sondes de la puce qui est insensible à une variation à courte distance mais se focalise plus sur des variations à longue distance. Une Waviness-SD élevée n'est pas forcément due à un excès de bruit de fond sur la puce mais peut être due au grand nombre d'altérations présent dans un échantillon. Ce critère est applicable pour des échantillons extraits de sang ou de lignées cellulaires, il n'est pas applicable à des échantillons tumoraux, pour ces échantillons il vaut mieux utiliser la **ndWavinessSD** définie dans le cadre de l'analyse **Normal diploïd ou de prélèvement FFPE en Oncoscan**: elle ne prend en compte que la variabilité des sondes dans les régions diploïdes de l'échantillon.

❖ Seuils recommandés des critères qualité en Single-sample analysis (puce Cytoscan HD):

- **MAPD $\leq 0,25$ pour Cytoscan**
- **SNPQC ≥ 15**
- **Waviness-SD $\leq 0,12$ (pour échantillons extraits de sang ou de lignées cellulaires)**

❖ Seuils recommandés des critères qualité en Normal diploïd analysis (puce Cytoscan HD et Oncoscan):

- **MAPD $< 0,25$ pour Cytoscan et $< 0,3$ pour Oncoscan**
- **ndSNPQC > 26**
- **ndWaviness-SD $< 0,12$**

Pour les puces Cytoscan ces critères ont été définis pour des échantillons dérivant du sang. Les critères SNPQC et Waviness-SD sont basés sur le postulat que les échantillons sont diploïdes et sans mosaïcisme ce qui est rarement le cas dans les échantillons tumoraux. Le fournisseur recommande que pour les échantillons tumoraux, **seul le critère MAPD soit considéré**. Si la dispersion des log2 ratio autour de la ligne de base est trop importante et donc que le MAPD est trop élevé le cas doit être rendu **non interprétable**.

ndCount : critère qualité dans le cadre de l'analyse Normal diploïd : cette valeur correspond au nombre de marqueurs diploïdes normaux identifiés pour faire l'analyse Normal diploïd. Si cette valeur descend sous un minimum l'alerte : « Low Diploïd Flag » indique "Low Diploid Flag = Yes". Cela signifie que le nombre de marqueurs est trop bas. Un ndCount $>10\,000$ est souhaité pour que le ndSNPQC et la centralisation soit considérée comme fiable.

Quand l'alerte : « Low Diploïd Flag » indique "**Low Diploid Flag = Yes**" :

- la ndWavinessSD est calculée alors sur tous les autosomes
- le ndSNPQC est calculé sur une préselection de 10 000 marqueurs autosomiques

6.2.3.1.2 Cas particulier Oncoscan CNV

➤ **CelPairCheck Status** : est un test qui vérifie chaque paire d'intensité entre les fichiers .cel (AT et GC) afin de déterminer si les fichiers ont été assemblés correctement et s'ils ont été assignés au bon canal (AT ou GC). Il vérifie les génotypes dans les deux fichiers d'intensité afin de voir si l'échantillon n'est pas le même entre les puces AT et GC sensées appartenir au même échantillon.

➤ **CelPairCheck Compare rate** : correspond au pourcentage de la signature SNP des marqueurs contrôles dont les génotypes sont comparés entre les canaux AT et GC. Ce taux doit être au-dessus d'une valeur minimum pour que ce test détermine si les fichiers AT et GC appartiennent bien au même individu.

➤ **CelPairCheck Concordance** : correspond à la concordance des génotypes d'un groupe de SNP des marqueurs contrôles comparés entre les fichiers AT et GC. Si le « compare rate » est haut mais que la concordance est faible le statut CelpairCheck indiquera un message d'erreur : « PossibleCELmispair ».

Si lors de ces vérifications faites par la console Oncoscan il y a un message d'erreur ne pas poursuivre l'analyse :

CelPairCheckStatus

CelPairCheckStatus message	Description
Pass	No intensity file pairing problem is detected.
PossibleCELmispair	Low genotype concordance exists between data in AT channel and data in GC channel. This is consistent with the data in the two intensity files originating from different individuals. This message may also appear with poor data. Please review the channel-specific sample QC metrics to see if one or both CEL files has a problem.
PossibleGCinATchannel	Data from the GC reaction appears to be assigned to the AT analysis channel.
PossibleATinGCchannel	Data from the AT reaction appears to be assigned to the GC analysis channel.
Warn	Cannot determine whether the two intensity files belong to the same individual. This message may appear with poor data, or with good data where too many of the signature SNPs used for CelPairCheck are in chromosome regions with copy number aberrations.

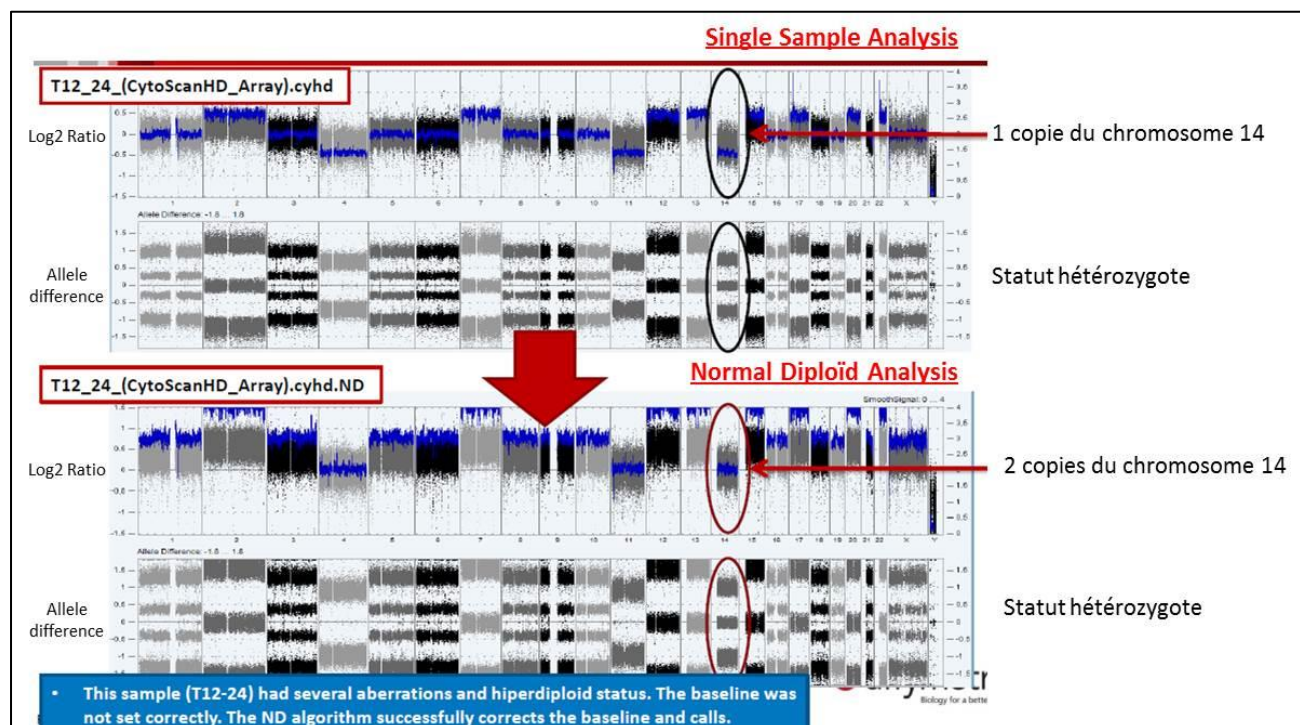
- **Ne pas analyser ou essayer de forcer les appariements** : en référer à un supérieur et retracer les erreurs possibles :
 - vérifier l'identité des tubes d'ADN initiaux et leur ordre d'utilisation
 - vérifier l'ordre des tubes conservés tout au long de la technique
- Tracer toutes les vérifications faites dans la feuille de série (**PAM-FQ-0181**) et indiquer l'origine du problème si elle a été identifiée

Faire une non-conformité ou une déclaration d'incident dans le SIL selon le problème (BPAT-PROC-0030) et voir la marche à suivre avec l'ingénieur, son supérieur et/ou le pathologiste demandeur selon l'erreur

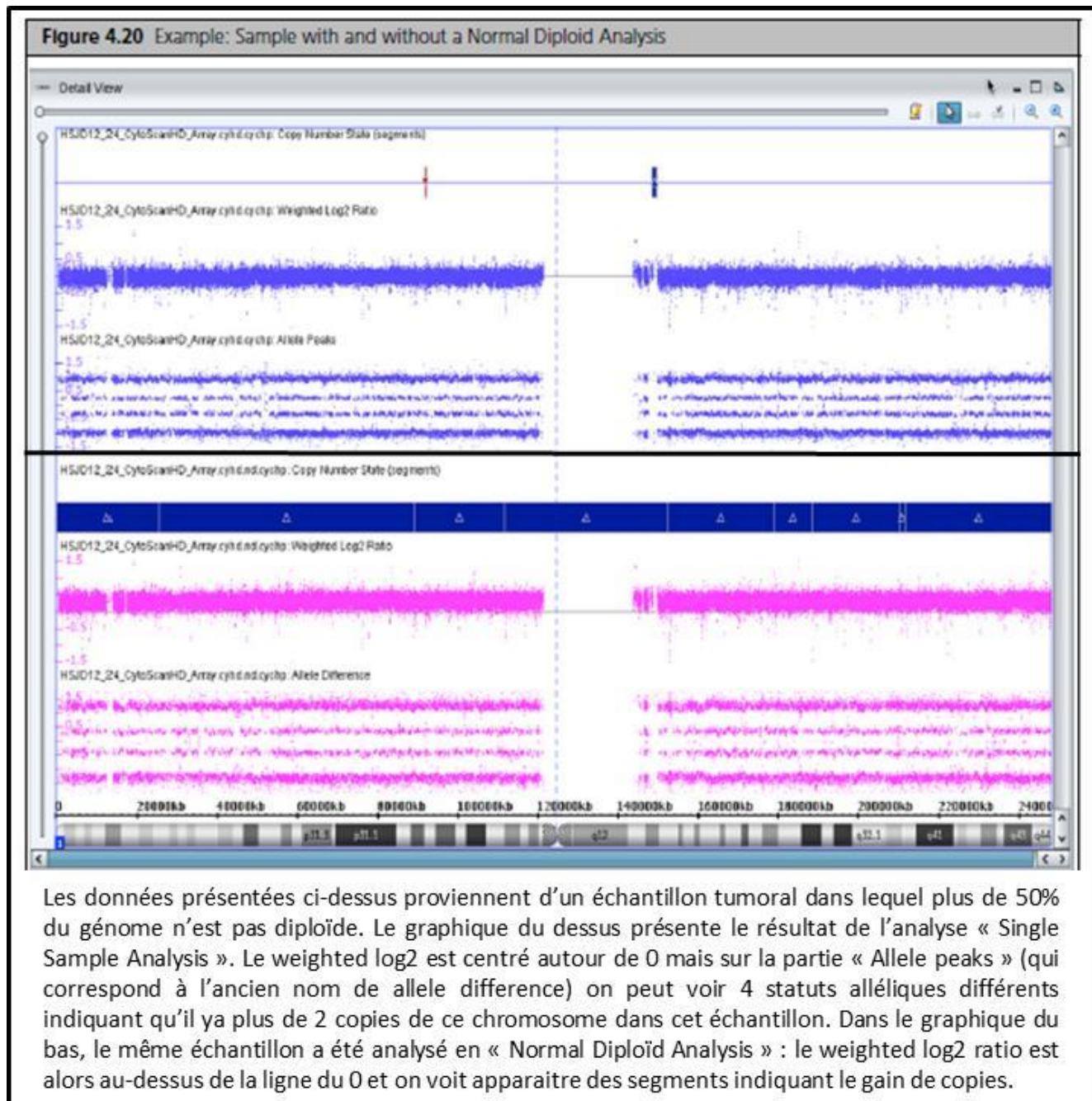
6.2.3.2 Centralisation

- ❖ Affymetrix propose deux types d'analyse pour les puces **Cytoscan HD**:
 - **The single-sample analysis:** l'intensité de fluorescence de la puce est calibrée par rapport au fichier de référence qui compile les résultats d'un millier d'échantillons
 - **The Normal diploid analysis:** l'analyse détermine automatiquement les régions normales diploïdes dans l'échantillon et normalise le reste de l'échantillon par rapport à ces régions permettant une centralisation adaptée des données. Cette analyse calcule le nombre de copies et vérifie si le profil allélique est correct. Si ce dernier n'est pas correct elle recalcule jusqu'à ce que ce soit le cas. **Pour pouvoir lancer cette analyse il faut au moins qu'une moitié de chromosome soit à 2 copies sans LOH dans l'échantillon.**

Un point à regarder en priorité est le **statut allélique des pertes de chromosomes ou de régions chromosomiques** : si une région est perdue en « Single Sample Analysis » avec un statut allélique hétérozygote il faut lancer la « Normal Diploid analysis » pour corriger la centralisation du cas :



Autre exemple de correction apportée par l'analyse « Normal Diploid analysis » :



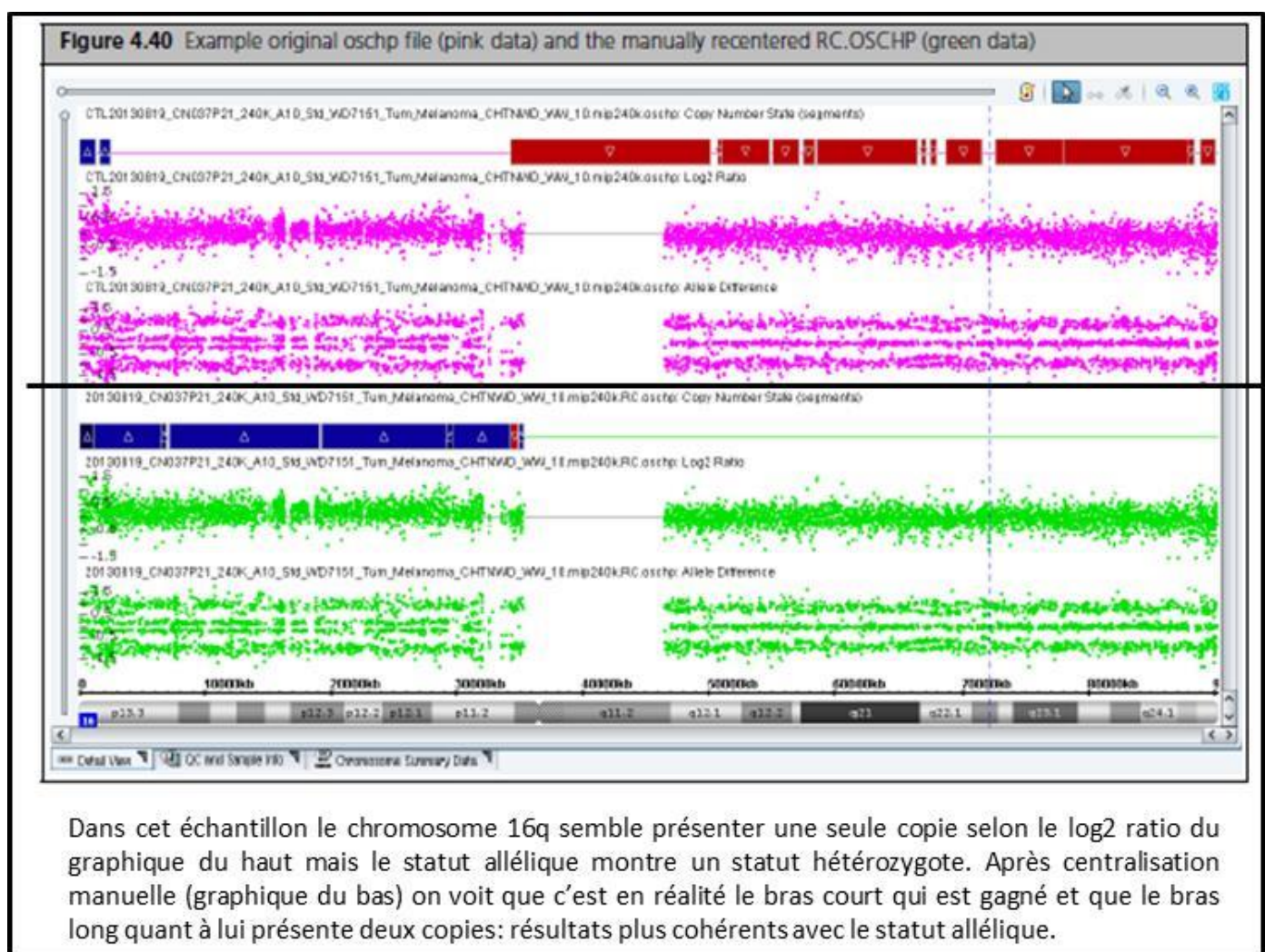
Le fournisseur recommande de lancer en première intention l'analyse Single Sample Analysis et s'il y a un doute concernant la ploïdie d'un échantillon alors l'analyse Normal Diploid Analysis peut être lancée pour cet échantillon (ne pas la lancer sur un échantillon normal car ça génère du bruit de fond).

Le fournisseur recommande de lancer l'analyse Normal Diploid pour des échantillons dans lesquels **plus de 50%** du génome est remanié.

❖ Affymetrix propose un seul type d'analyse pour les puces **Oncoscan CNV**:

Pour les puces **OncoScan CNV** dans la grande majorité des cas il n'y a pas besoin de changer d'algorithme d'analyse car OncoScan est une puce conçue pour le cancer et donc les ploïdies complexes : l'algorithme calcule automatiquement et vérifie la cohérence entre le copy number et le profil d'allèles. Cytoscan est à la base une puce cytogénétique et donc conçue pour des ploïdies simples. Pour le cancer, l'algorithme Oncoscan a été adapté à la CytoScan avec l'analyse « Normal Diploïd ».

Toutefois dans certains cas comme dans l'exemple ci-dessous il peut être nécessaire de recentrer l'échantillon, pour cela se référer aux documents **PAM-DEXT-0131** (pages 63 à 67) et **PAM-MOP-0075**:



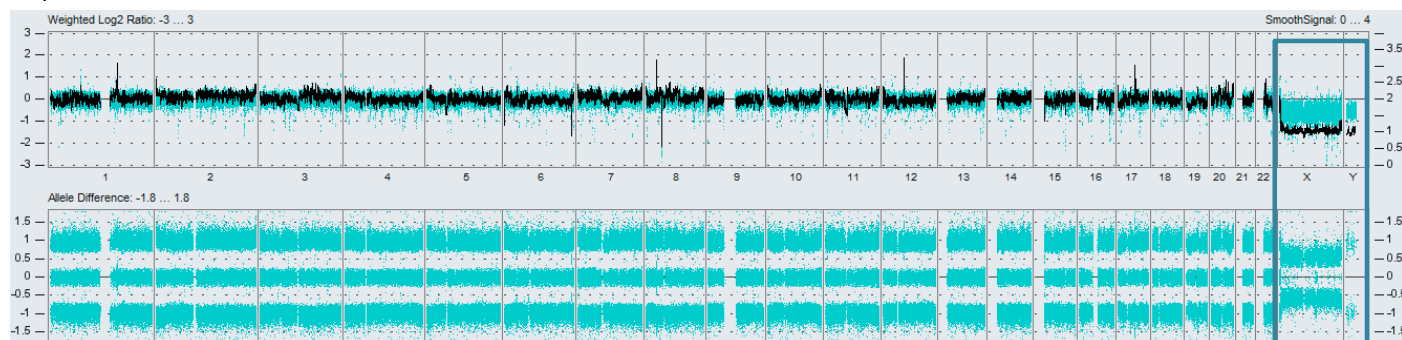
6.2.3.3 Vérification du sexe obtenu en CGH

Le logiciel définit le : **Y Gender Call** : le sexe de l'échantillon est défini par rapport au signal sur le chromosome Y.

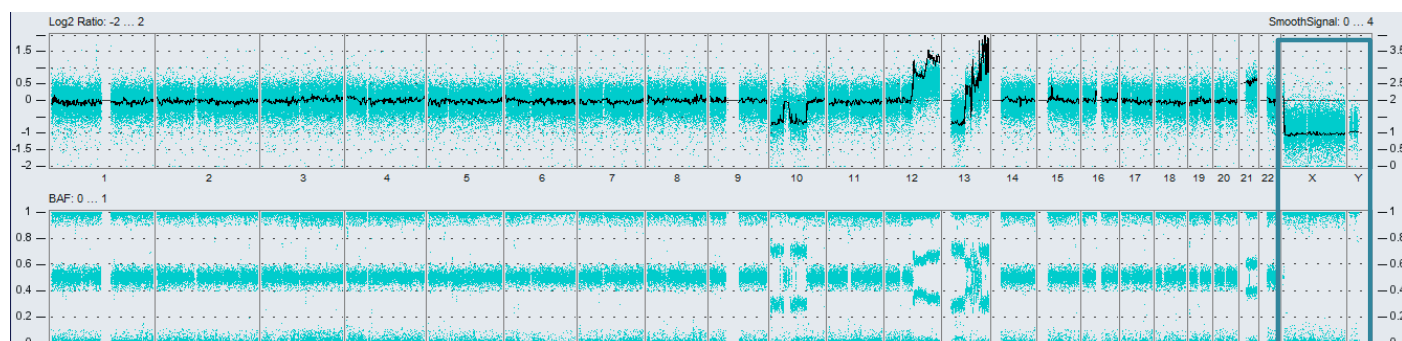
Toutefois il est nécessaire de le vérifier soi-même en regardant les profils obtenus sur les chromosomes X et Y :

- ✓ La tumeur doit présenter si c'est celle d'un **homme** : 1 copie avec un statut allélique homozygote pour les chromosomes X et Y

Cytoscan HD: homme

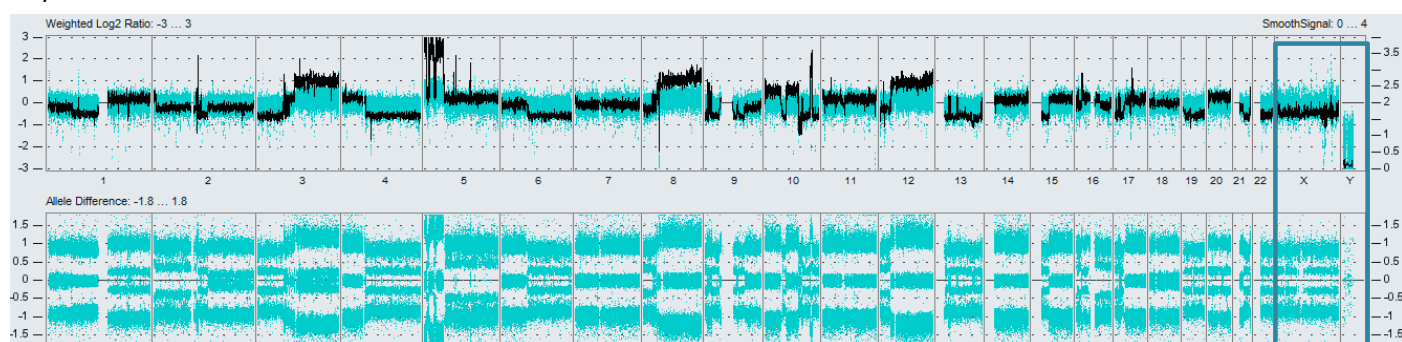


Oncoscan CNV: homme

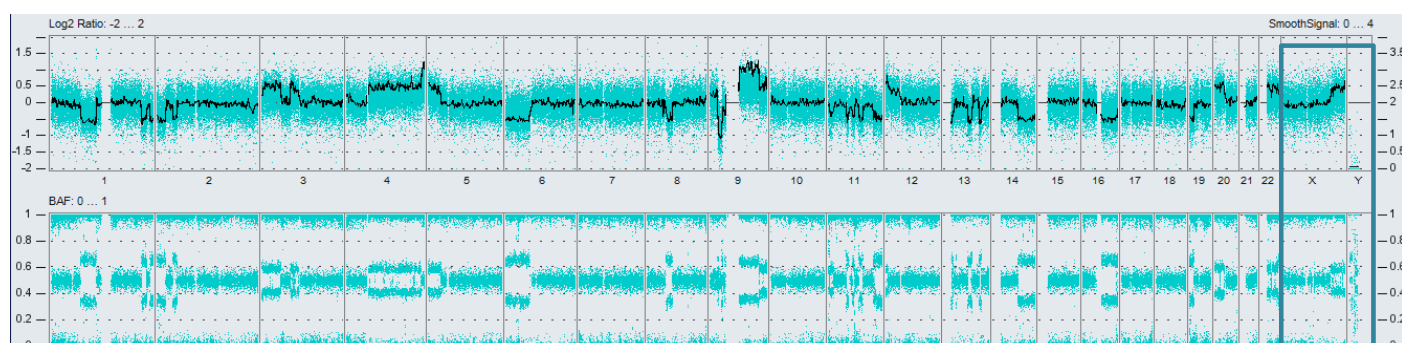


- ✓ La tumeur doit présenter si c'est celle d'une **femme** : 2 copies avec un statut allélique hétérozygote pour le chromosome X et aucune copie du chromosome Y

Cytoscan HD: femme



Oncoscan CNV: femme



Attention rester conscient qu'une femme ne peut pas avoir de chromosome Y mais par contre la tumeur d'une femme peut perdre un chromosome X tout comme la tumeur d'un homme peut perdre un chromosome Y ou gagner un ou plusieurs chromosomes X. Toujours vérifier tous les sexes de tous les échantillons à la suite et interpréter la série dans sa globalité.

S'il y a un problème de cohérence entre le sexe obtenu en CGH et le sexe du patient, en référer à un supérieur et :

- vérifier s'il n'y a pas eu d'erreur de saisi du sexe dans le SIL à réception des prélèvements
 - vérifier l'identité des tubes d'ADN initiaux et leur ordre d'utilisation
 - vérifier l'ordre des tubes conservés tout au long de la technique
 - faire une PCR SRY sur les tubes d'ADN initiaux pour vérifier le sexe de l'échantillon extrait afin de voir si le problème pourrait venir de la technique de CGH ou s'il est antérieur (**PAM-FQ-0170**)
-
- Tracer toutes les vérifications faites dans les feuilles de série (**PAM-FQ-0171**, **PAM-FQ-0180** ou **PAM-FQ-0181**) et indiquer l'origine du problème si elle a été identifiée
 - Faire une non-conformité ou une déclaration d'incident dans le SIL selon le problème, avertir immédiatement le pathologiste demandeur et voir la marche à suivre avec l'ingénieur, un médecin de PAM et/ou le pathologiste demandeur selon l'erreur (**BPAT-PROC-0030**)

6.2.3.4 Effet bandes R

Vérifier l'absence d'effet bandes R sur le profil (confère paragraphe **6.1.2.4**)

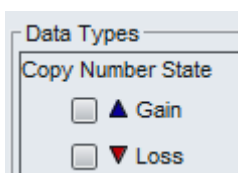
6.2.4 Analyse

- L'analyse ne doit pas être réalisée si les vérifications préalables n'ont pas été faites par le technicien et le médecin de manière indépendante et s'il y a le moindre doute quant à un des points mentionnés dans la partie au-dessus
- Le manuel fournisseur concernant le logiciel ChAS est le suivant : **PAM-DEXT-0130**
- Les documents fournisseurs d'exemples d'analyse des puces à ADN sont les suivants :

- **PAM-DEXT-0131**: fourni des exemples d'analyse de puces Oncoscan CNV
- **PAM-DEXT-0132**: permet de comprendre la représentation en différence allélique
- **PAM-DEXT-0133**: fourni des exemples d'analyse de puces Cytoscan HD

- Quelle que soit le type de puce il est possible de visualiser les informations suivantes :

- Le **log2 ratio** qui correspond à l'intensité pour chaque sonde « Z » normalisée par rapport à la référence. Lorsque la sonde est à 2 copies dans l'échantillon le graphe est centré autour du 0 sur l'axe Y ($\log_2(1)=0$).
- Le **Weighted log2 ratio** : correspond au log2 ratio corrigé par "Bayes wavelet shrinkage estimator"
- Le **Copy number state** : l'algorithme évalue le nombre de copies correspondant au log2 ratio de la région
- **LOH** : le logiciel indique les segments présentant une perte d'hétérozygotie par un symbole
- **L'allèle différence** : chaque allèle a une valeur de 0.5 et c'est la différence entre l'allèle A – l'allèle B qui est représentée
- **LOH** : met en évidence les régions de perte d'hétérozygotie
- **Smooth signal** : le signal lissé estimant le nombre de copies
- **BAF** : B allele frequency, représente la proportion du signal expliqué par un seul allèle, l'allèle B
- **Gain et loss** : le logiciel peut afficher les régions gagnées (▲) ou perdues (▼) par segment si dans la partie « Data types » gain et loss sont sélectionnés :



- **GainMosaic et LossMosaic** : affichage spécifique aux puces Cytoscan (cf paragraphe 6.2.4.3)

Cytoscan

Data Types	Definition
Detected Segment Types	
GainMosaic	Non-integer amplifications or duplications
LossMosaic	Non-integer hemizygous or homozygous deletions
Gain	Amplifications or duplications
Loss	Hemizygous or homozygous deletions
LOH	Loss of Heterozygosity
Probe array data (displayed as graph data)	
Copy Number State	HMM-derived integer Copy Number State
Log2 Ratio	Per marker Log2 Ratio of normalized intensity with respect to a reference, with further correction for sample specific variation.
Weighted Log2 Ratio	Contains the Log2 Ratios processed through a Bayes wavelet shrinkage estimator. These processed values are input to the CNState algorithm HMM.
LOH	Loss of Heterozygosity
Allele Difference	Filtered and smoothed values for individual markers. Nonparametric estimation is used to understand possible regional peak structure towards which the data is smoothed. The amount of filtration and smoothing is dynamically adapted based on sample quality. Allele difference is computed based on differencing A signal and B signal, then standardizing based on reference file information.
Genotype Calls	SNP genotype calls (single sample, BRLMM-P-plus algorithm)
Smooth Signal	Gaussian Smoothed Calibrated Copy Number Estimate
B-allele Frequency	Number of B alleles/number of A+B alleles used to show allelic imbalances.

Oncoscan

Data Types	Definition
Detected Segment Types	
Gain	Amplifications or duplication
Loss	Hemizygous or homozygous deletions
LOH	Loss of Heterozygosity
Probe array data (displayed as graph data)	
Copy Number State	TuScan derived Copy Number State.
Log2 Ratio	Per marker Log2 Ratio of normalized intensity with respect to a reference, with further correction for sample specific variation.
Weighted Log2 Ratio	Contains the Log2 Ratios processed through a Bayes wavelet shrinkage estimator. These processed values are input to the CNState algorithm HMM.
LOH	Loss of Heterozygosity.
Allele Difference	Difference of A signal and B signal, each standardized with respect to their median values in the reference.
Smooth Signal	Smoothed Calibrated Copy Number Estimate.
Somatic Mutations	Location and detection of Somatic Mutation. (OncoScan FFPE Assay only)
B-allele Frequency	Number of B alleles/ number of A+B alleles, used to show allelic imbalances.

PAM-DEXT-0130

6.2.4.1 Log2 ratio

- L'analyse ne doit pas être réalisée si les vérifications préalables n'ont pas été faites par le technicien et le médecin de manière indépendante et s'il y a le moindre doute quant à un des points mentionnés dans la partie au-dessus
- **Analyse : toujours interpréter dans le contexte d'une région donnée, une seule sonde altérée peut-être artefactuelle et toujours interpréter en considérant le statut allélique et en confrontant le résultat avec l'interprétation du logiciel**
 - ✓ Un statut **normal** correspond à un log2 ratio compris entre -0,25 et 0,25
 - ✓ Un **gain** correspond à un log2 ratio compris entre 0,25 et 2
 - ✓ Une **amplification** correspond à un log2 ratio supérieur à 2
 - ✓ Une **perte hétérozygote** correspond à un log2 ratio compris entre -0,25 et -1
 - ✓ Une **perte homozygote** est suspectée quand le log2 est inférieur à -1

Attention pour des altérations de quelques sondes sur une région il faut vérifier qu'il ne s'agisse pas d'une région de type CNV (copy number variation) qui sont des polymorphismes de nombre de copies et qui donc peuvent différer entre le patient et l'ADN de référence utilisé

6.2.4.2 B allele frequency (BAF)

Pour obtenir le génotype d'une région il faut **toujours** interpréter à la fois le BAF et le copy number (CN) et/ou le log2 ratio.

Il faut d'abord regarder le CN ou log2 ratio puis interpréter le BAF selon ce nombre de copies car un même BAF peut avoir une interprétation différente selon le CN (mosaïcisme ou gain ...).

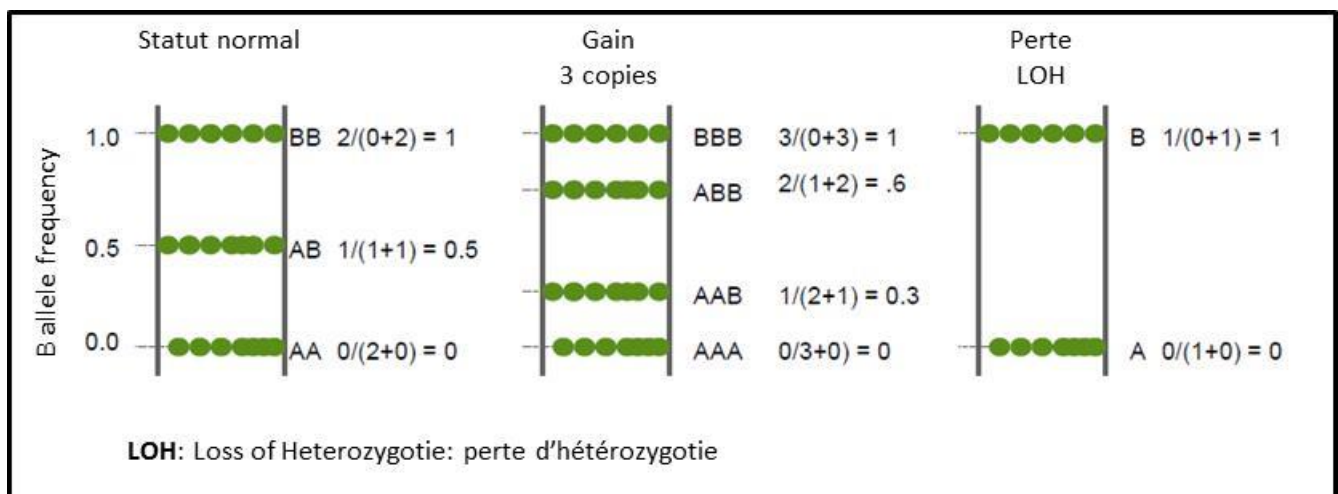
Le BAF représente la proportion du signal expliqué par un seul allèle, l'allèle B.

Les valeurs du BAF peuvent aller de -1 à 1 et il est calculé selon la formule : $B/(A+B)$:

Pour une sonde donnée :

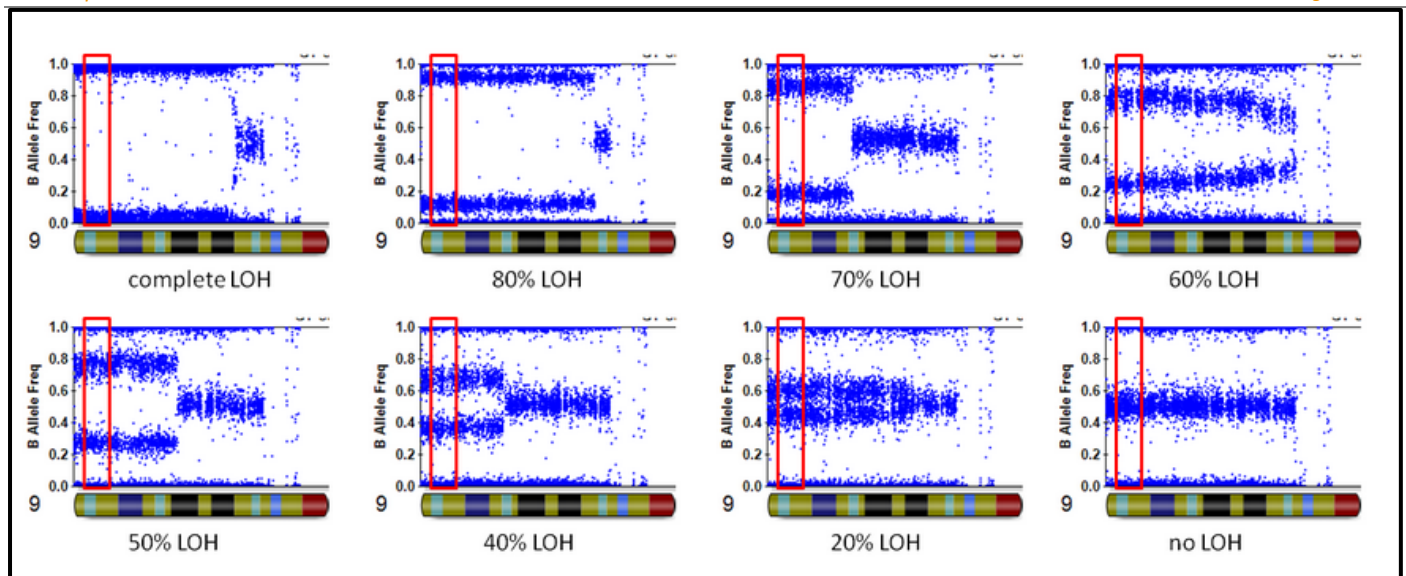
- Si le **BAF =0**, cela signifie qu'il n'y a pas d'allèle B donc que selon le CN le génotype est AA ou A ou AAA ou AAAA...
- Si le **BAF =0,5**, le statut est hétérozygote et selon le CN le génotype peut être AB, AABB ...
- Si le **BAF =1**, cela signifie qu'il n'y a pas d'allèle A donc que selon le CN le génotype est BB ou B ou BBB ou BBBB...
- Si le **BAF =0,66**, cela signifie qu'il y a plus de B que de A donc le statut peut être par exemple ABB...

Pour une région donnée : (PAM-DEXT-0131)



Mosaïcisme :

Il faut également prendre en compte que la tumeur peut être hétérogène et/ou peut être contaminée par des cellules normales ce qui va faire varier les valeurs du BAF : plus la contamination ou l'hétérogénéité est grande plus les valeurs du BAF vont tendre vers 0 (exemple pour une région à 2 copies) :



Pour plus d'exemples de mosaïcisme selon le nombre de copies de la région se référer au document **PAM-DEXT-0131** pour les puces **Oncoscan** et **PAM-DEXT-0133** pour les puces **Cytoscan**.

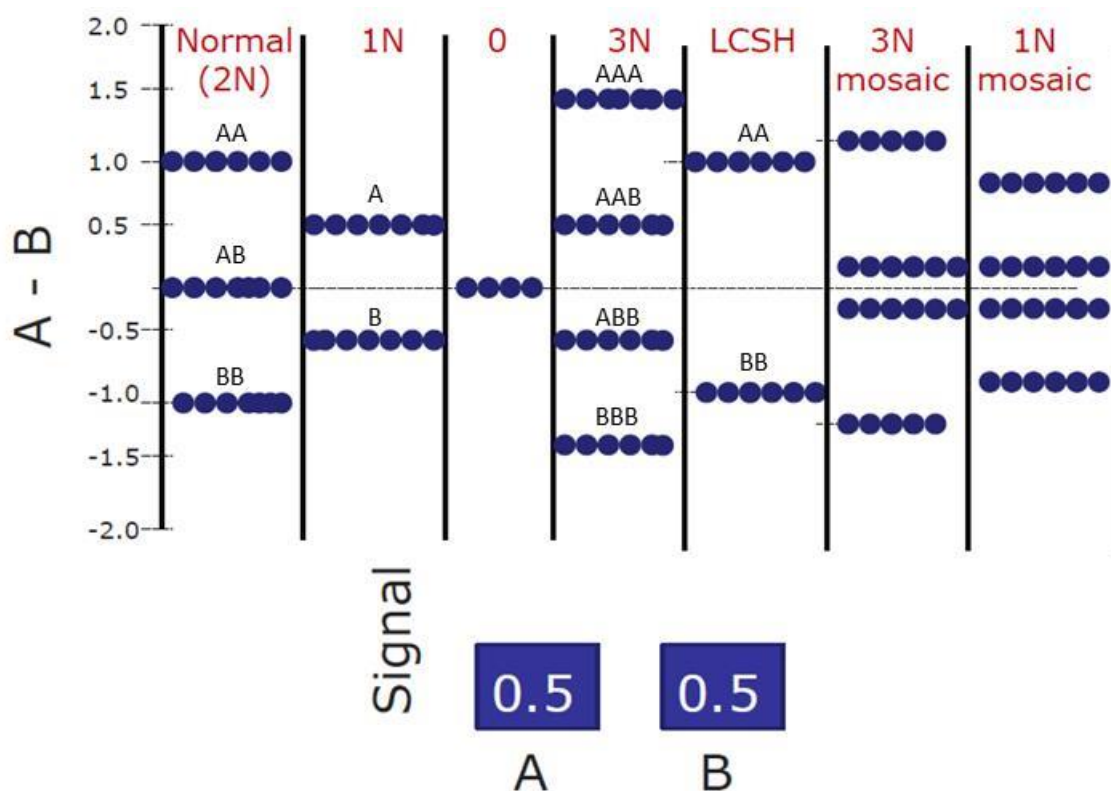
6.2.4.3 Allele difference

Une autre représentation du statut allélique est disponible dans ChAS : l'allèle différence.

Pour chaque SNP il y a deux possibilités : l'allèle A ou l'allèle B (chaque allèle est mesuré par une sonde spécifique). Dans cette représentation un allèle A et un allèle B contribuent chacun pour 0,5 unité du statut allélique et dans la représentation de l'allèle différence la valeur globale de B est soustraite à la valeur de A (A-B). S'il y a deux allèles A par exemple et deux allèles B (CN=4 et génotype = AABB), la valeur globale des A est 0,5+0,5 donc 1 et de la même manière pour la valeur globale de B donc A-B= 0.

Pour une sonde donnée :

- Si le **A-B =0**, cela signifie qu'il y a autant de copies de A que de B donc selon le CN le génotype est AB ou AABB ...
- Si le **A-B =0,5**, cela signifie qu'il y a plus de A que de B, le statut allélique peut donc être selon le CN : A, AAB ...
- Si le **A-B =-0,5**, cela signifie qu'il y a plus de B que de A, le statut allélique peut donc être selon le CN : B, ABB ...
- Si le **A-B =1**, cela signifie qu'il n'y a pas d'allèle B si le CN=2 ou qu'il y a plus de A que de B si le CN=4 par exemple donc que selon le CN le génotype peut être AA ou AAAB ...
- Si le **A-B =-1**, cela signifie qu'il n'y a pas d'allèle A si le CN=2 ou qu'il y a plus de B que de A si le CN=4 par exemple donc que selon le CN le génotype peut être BB ou AB BB ...
- Si le **A-B** a une valeur non entière différente d'un multiple de 0,5 cela signifie qu'il y a un mosaïcisme

Pour une région donnée : (PAM-DEXT-0132) exemples de résultats

6.2.4.4 Pourcentage de cellules aberrantes : mosaïcisme

Le mosaïcisme est suspecté quand :

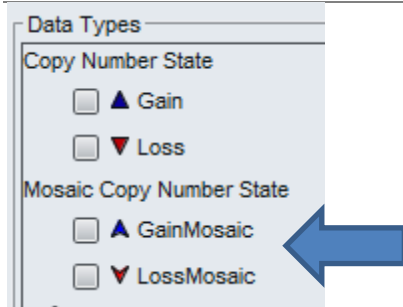
- Le smooth signal présente pour une région donnée une valeur non entière (exemple 2,9 copies)
- L'allèle différence a une valeur non entière d'un multiple de 0,5

6.2.4.4.1 Puce Cytoscan

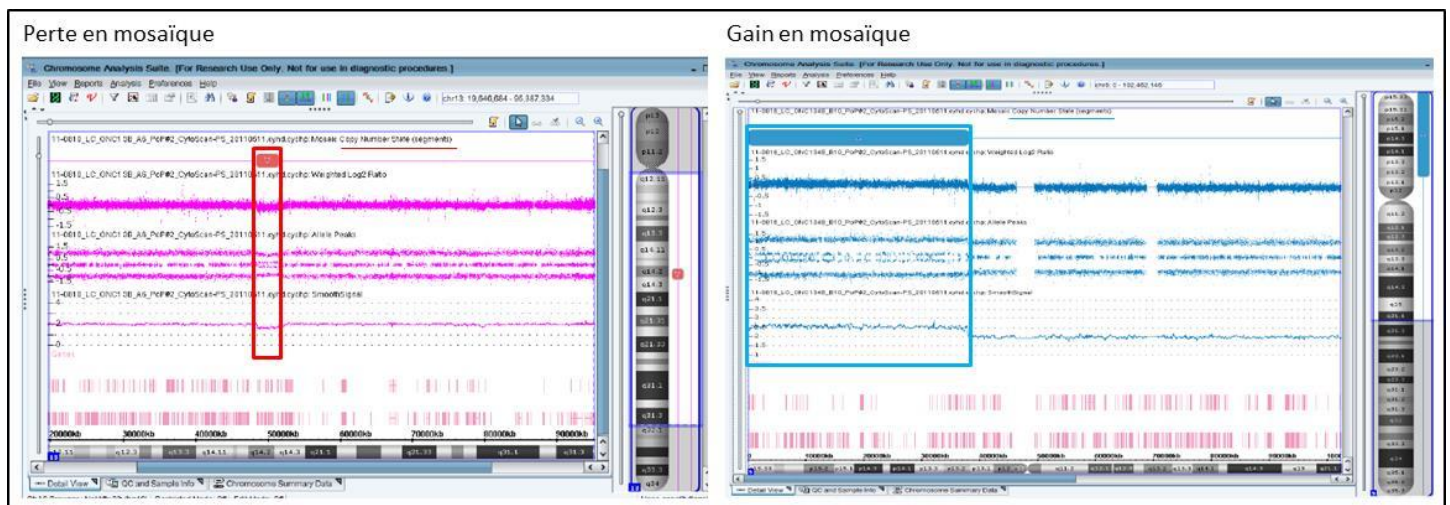
- Pour les puces **Cytoscan HD** le logiciel peut mettre en évidence les régions présentant soit un gain soit une perte en mosaïque.

Toutefois le logiciel ne détecte que les mosaïques entre 30 et 70% pour un CN entre 1 et 3 sur des régions d'au moins 5000 sondes. L'algorithme ne considère que les niveaux de mosaïque de 30, 50 et 70%. Pour détecter un mosaïcisme il est défini que le range des log ratios doit être cassé en une série de bandes selon le niveau de détection (30% ou plus, 50% ou plus et 70 à 100%). Les mosaïques pour le chromosome Y ne sont jamais détectées.

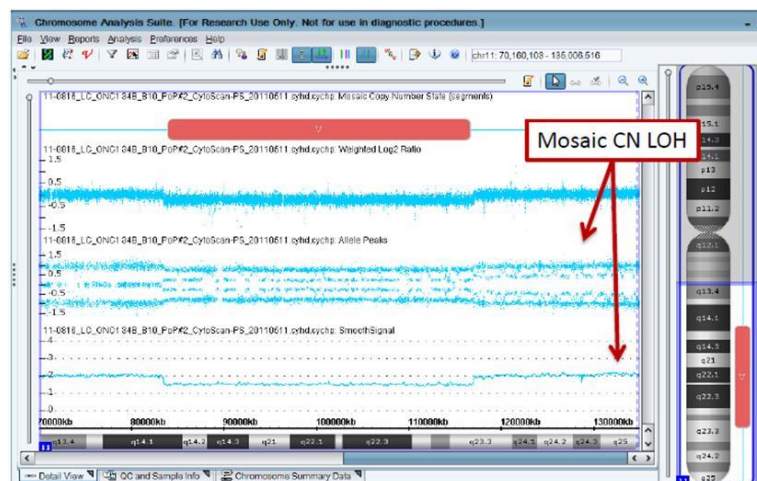
- Pour visualiser les mosaïques détectées par le logiciel, il faut sélectionner dans la partie « Data Types » : GainMosaic et LossMosaic :



- Ainsi le logiciel affichera dans la partie où sont affichées toutes les représentations pour la région visualisée une ligne supplémentaire avec les segments : Mosaic Copy Number State (segments) : en bleu les gain avec le symbole ▲ et en rouge les perte avec le symbole ▼ :



- Mais attention pour les LOH en mosaïque avec un nombre de copie neutre (CN=2 par exemple) le logiciel ne mettra pas en évidence la mosaïque :



Pourcentage de cellules aberrantes

- L'algorithme Tuscan utilise le BAF et le log2 ratio pour estimer la ploïdie et le pourcentage de cellules aberrantes dans l'échantillon (%AC) qui a son tour est utilisé pour calculer le CN. Le BAF et le log2 ratio contribuent à part égale pour le calcul du CN.

- Tuscan les utilise d'abord pour définir les segments avec le même nombre de copies. Ensuite il utilise le BAF, le log2 ratio et les données de segmentation pour définir la combinaison de % AC et la ploïdie qui correspondent le mieux avec les données.
- Quand Tuscan arrive à définir le % AC il assigne à chaque segment aberrant un CN entier qui représente le CN de la portion tumoral de l'échantillon. Cela est possible quand la tumeur est assez homogène.
- Si la tumeur est très hétérogène (pas de clone dominant) ou qu'elle contient beaucoup de cellules normales le logiciel n'arrive pas à définir le % AC. En d'autres termes si le %AC contribuant aux différentes aberrations varie selon les aberrations, la ploïdie et le % AC ne peuvent être définis.
- Quand le % AC n'est pas défini, l'algorithme de segmentation identifie quand même les segments avec le même CN mais le CN des cellules aberrantes ne peut pas être défini. Dans ce cas, Tuscan attribue des valeurs de CN fractionnées en les incrémentant d'un tiers (2.33, 2.66, 3 etc).
- **Ainsi les utilisateurs doivent regarder si le CN est celui des cellules aberrantes (% AC=un chiffre) ou s'il représente l'intégralité de l'échantillon (cellules normales et tumorales ou mélange de cellules tumorales différentes : % AC=NA).**
- L'hétérogénéité de la tumeur affecte également l'interprétation du CN. En effet, par exemple un CN de 2.33 peut être dû à 40% de cellules aberrantes présentant 3 copies ou bien 10% de cellules présentant 5 copies ou un mélange encore plus complexe d'hétérogénéité.

6.2.4.4.2 Puce Oncoscan CNV

- Pour les puces **Oncoscan CNV** le logiciel ne permet pas d'afficher les GainMosaic ou les LossMosaic.

6.2.4.5 Retour d'hétérozygotie

On peut suspecter l'existence d'une réelle délétion homozygote quand on a vérifié d'une part que le profil était bien centralisé, que le CN pour une région indique 0 copie et qu'on a en statut allélique des sondes présentant un statut hétérozygote à 0 qui traduit en fait le statut des cellules normales contaminantes.

6.3 Rendu de résultat

- Le praticien qui a analysé le profil génomique fait une synthèse analytique dans le SIL (**PAM-MOP-0044**), il décrit les altérations observées et donne une conclusion quant à la question posée par le pathologiste
- Le praticien demandeur de l'examen fait un CR de Pathologie Moléculaire à la suite de la synthèse analytique et complète éventuellement ensuite son compte-rendu d'anatomopathologie si besoin (**PAM-PROC-0017**)