Copy Number and Genotype Analysis of FFPE-extracted DNA
Recommendations and Guidelines for GeneChip® Mapping Arrays

Introduction

Numerous publications demonstrate the value of using GeneChip® Mapping Arrays to provide genome-wide analysis of genotype, copy number and loss of heterozygosity (LOH) from high-quality DNA obtained from blood or cell lines. The standard practice in pathology labs, however, is to collect and store clinical tissue as formalin-fixed, paraffin-embedded (FFPE) samples. Unfortunately, the FFPE process often yields DNA that is degraded and/or contaminated with chemical residues which interfere with standard molecular techniques. Variability in processing between institutes and users, DNA extraction methods and tissue source or time of storage can also influence the quality of DNA from an FFPE sample.

The following technical note provides guidelines for sample and data processing to ensure the best results from 500K analysis of FFPE samples. It includes recommendations for paraffin extraction, QC testing of input DNA, selection of the most appropriate mapping array product and data analysis options.
**Recommended Workflow**

1. **EXTRACT DNA FROM FFPE SAMPLES**

Several high-quality commercial kits are currently on the market for extracting DNA from FFPE samples. We have successfully extracted DNA of sufficient quality for the mapping arrays using the Qiagen DNeasy Blood & Tissue Kit®, with modifications. Suggested modifications to this protocol include:

- 95°C, 15-minute treatment prior to Proteinase K digestion, in ATL buffer
- Three days of Proteinase K treatment at 56°C (with daily spike-ins of the same amount of PK used on the first day)
- NH4OAc + EtOH cleanup of extracted DNA (as described as an optional cleanup step in the Mapping Assay Manual). Other extraction protocols were evaluated and found to provide FFPE DNA of reasonable quality for application to the mapping arrays (by containing sufficient fragment size and exhibiting sufficient performance during PCR). These included Argylla Technologies PrepParticles (beta version) (http://argylla.com/)

2. **QC INPUT DNA USING PCR-BASED TEST**

Not all FFPE DNA samples are alike. Differences in fixation protocols, length of storage times, tissue source and extraction methods may affect the quality of DNA that is produced from an FFPE sample. Therefore, it is essential that each FFPE DNA sample be tested for quality prior to applying a sample to the mapping array.

A PCR-based test should be used to determine the quality of an FFPE DNA sample. This QC step allows the user to ask two important questions:

- Is the DNA amenable to PCR amplification?
- What size fragments can I amplify from this DNA sample?

The answers to these questions will be influenced by both the extent of degradation in the sample and by the degree of inhibition by contamination or chemical residues.

There are several PCR-based tests that can be applied to an FFPE sample. Possibilities include multiplex PCR, RAPD PCR (Siwoski, et al., 2002), or the digestion/ligation/PCR steps in the mapping assay (for more information, see Table 1 and Figure 1). These QC examples all include the amplification of a variety of fragment sizes during PCR, including larger fragment sizes above 800 bp. Samples that can be applied to the array should provide large enough PCR products that a significant number of SNPs can be used for genotype and copy number analysis. Table 2 and Table 3 indicate the number of SNPs represented on various fragment size groups.

![Figure 1: Examples of FFPE samples that passed PCR-based QC tests (above) and failed the PCR-based QC test (below). DF = dilution factor; + = nonFFPE template.](image)

<table>
<thead>
<tr>
<th>RAPD PCR</th>
<th>Multiplex PCR</th>
<th>Mapping PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>DF: 1 10</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>600 bp</td>
<td>300 bp</td>
<td>600 bp</td>
</tr>
<tr>
<td>800 bp</td>
<td>300 bp</td>
<td>300 bp</td>
</tr>
</tbody>
</table>

![Table 1: Possible PCR-based pre-screening QC methods.](image)

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Multiplex PCR</strong></td>
<td>Amplification of several large amplicons (e.g. a range of 200 to 800 bp)</td>
<td>Easiest to implement</td>
</tr>
<tr>
<td><strong>RAPD PCR</strong></td>
<td>Use of 10-mer primers to simultaneously amplify multiple regions of genomic DNA, up to ~ 2 kb</td>
<td>See Siwoski, et al. This is a protocol, not a kit. The following details are suggested: (1) Use Qiagen HotStar Taq, &amp; (2) Use EtBr to visualize products on an agarose gel.</td>
</tr>
<tr>
<td><strong>Mapping PCR</strong></td>
<td>Digestion, ligation and PCR step from WGSA</td>
<td>Best predictor but uses most time, cost and DNA</td>
</tr>
</tbody>
</table>
The number of SNPs accessible is useful in determining whether or not to proceed with a particular sample. At a minimum, we do not recommend proceeding with samples that fail to produce PCR fragments greater than 300 bp.

3. SELECT APPROPRIATE MAPPING ARRAY

The mapping assay uses a technology termed Whole-Genome Sampling Analysis (WGSA), in which DNA is digested and ligated to adaptors that enable PCR amplification by one universal primer (Kennedy, et al., 2003). The PCR product is fragmented, labeled and hybridized to the arrays. All Affymetrix Mapping Arrays use WGSA for genotyping, and all support copy number analysis, but the Mapping 500K Arrays (including the Mapping 250K Nsp Array and the Mapping 250K Sty Array) are best suited for FFPE DNA analysis for the two reasons described below.

First, the size range of amplicons required by the PCR step during WGSA varies between arrays and is a limiting factor for degraded DNA samples such as FFPE DNA. The overall distribution of amplicon sizes is similar between the Mapping 10K Array and the Mapping 500K Array Set, with a maximum size of ~1,100 bp. In contrast, Mapping 100K Arrays rely on larger amplicons reaching 2,000 bp. For this reason, Mapping 100K Arrays are not recommended for application of FFPE DNA.

Second, while both Mapping 500K Arrays and the Mapping 10K Arrays are compatible with FFPE DNA, Mapping 500K Arrays offer more extensive coverage across the genome and are best suited to accommodate filters that exclude SNPs on large fragment size (see suggestions for data analysis below, and Figure 3).

Table 2 and Table 3 address these two main differences by listing the percentage of SNPs on each array within each fragment size group as well as the total number of SNPs per array. Although there are fewer total SNPs on the 250K Sty Array than the 250K Nsp Array, there are more SNPs on smaller fragments; therefore, the Mapping 250K Sty Array may be preferred when performing a single array experiment on degraded DNA. In summary, both the Mapping 500K Arrays and the Mapping 10K Arrays are suitable for analysis of FFPE DNA, but the Mapping 500K Arrays offer the best choice for this application.

4. PREPARE TARGET PER STANDARD WGSA PROTOCOL

We recommend following the standard WGSA protocol provided for the particular array selected. Special attention should be made to the following sections:

DNA Quantitation: Because there are often contaminants in an FFPE DNA sample that may affect optical density, we recommend...
quantitating FFPE DNA using a method that is specific to dsDNA concentration instead of using UV spectroscopy; for example, PicoGreen®.

Digestion, Ligation and PCR: No changes are recommended during the digestion and ligation steps of WGSAX. When setting up for PCR, the traditional protocol requires three PCR reactions per DNA sample, using a total of 30 µl from the 100 µl ligation product. The three PCR reactions are pooled together, and 90 µg of PCR product are carried over to the step. Because degraded FFPE DNA does not produce larger amplicons during PCR (Figure 1), the yield of PCR products will often be lower for the FFPE samples. Pooling additional PCR reactions to gather 90 µg of PCR product can increase call rates. Therefore, when setting up the PCR reactions, it is recommended to set up additional PCR reactions per FFPE DNA sample. Often, six total PCR reactions will be sufficient, but for severely degraded samples, nine may be required. Importantly, this step does not affect the input amount of DNA, which remains at 250 ng per array.

Purifying and Pooling PCR Reactions: For the Mapping 500K assay, Clontech® 96-well plates are used to pool and purify three PCR reactions per well. When more than three PCR reactions are performed for a single DNA sample, multiple wells should be used for purification. Pool up to three PCR reactions together in a single well, so that if there are six PCR reactions from one DNA source, there will be two wells used for PCR purification for this sample. After washing both wells three times with water and allowing the DNA to completely dry on the plate, elute the first well with 45 µl of RB buffer but leave the second well dry. Next, use the same RB buffer that now contains the DNA eluted in the first well to elute the DNA in the second well. In this way, all of the DNA will be eluted into the same 45 µl of RB.

Fragmentation and Labeling: There are currently no specific recommendations for altering the fragmentation and labeling steps. Use 90 µg when possible, but if less DNA was attained from PCR, it is okay to proceed with less. Decreased performance (number of SNPs providing reliable genotype and copy number data) should be expected with decreased DNA input at this step.

5. MODIFY DATA ANALYSIS TO ACCOUNT FOR DNA DEGRADATION

Genotypes: Genotypes should be assigned using the DM algorithm, an option in Affymetrix’ genotyping software GTYPE 4.1. Unlike traditional assessments of high-quality DNA samples, the overall call rate is not an indication of successful performance on the mapping array when using FFPE DNA; expect that genotyping call rates will be reduced for these samples. Many FFPE samples have provided call rates in the range of 70 to 95 percent.

SNPs on fragments that were not amplified robustly during PCR will have increased no-calls and decreased reliability. For dependable genotypes, exclude SNPs on the larger fragments. Fragment sizes can be attained in the CSV file associated with the mapping array, which can be found at: http://www.affymetrix.com/support/technica l/byproduct.affx?product=500k.

Loss of Heterozygosity: Because of high reliability of genotype calls for SNPs on smaller fragment sizes when using FFPE DNA, it is possible to analyze FFPE samples for loss of heterozygosity (LOH). Since genotyping reli-

<table>
<thead>
<tr>
<th>Table 3: SNP coverage across the Mapping 500K Arrays based on fragment size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragment Size Cutoff</td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>≤ 200 bp</td>
</tr>
<tr>
<td>≤ 300 bp</td>
</tr>
<tr>
<td>≤ 400 bp</td>
</tr>
<tr>
<td>≤ 500 bp</td>
</tr>
<tr>
<td>≤ 600 bp</td>
</tr>
<tr>
<td>≤ 700 bp</td>
</tr>
<tr>
<td>≤ 800 bp</td>
</tr>
<tr>
<td>≤ 900 bp</td>
</tr>
<tr>
<td>≤ 1kb</td>
</tr>
<tr>
<td>NO CUTOFF</td>
</tr>
</tbody>
</table>

Figure 3: Brown bars represent SNP locations from the Mapping 500K Array Set across a segment of chromosome 1. Various filters based on fragment size were applied to demonstrate the effect of these filters on genomic coverage. Blue regions represent genes from the RefSeq database that reside either on the sense (+) or the antisense (-) strand. In view are SNPs from both Mapping 500K Arrays.
ability of SNPs on larger fragments is reduced for degraded DNA samples, we recommend excluding SNPs on larger fragment sizes when assessing genotype calls for LOH. This is an option in most software packages that provide LOH estimates from Affymetrix Mapping Array data.

Copy Number: Software choice. Most copy number analysis tools do not correct for the effects of DNA degradation and contamination of FFPE samples. Currently, only two software solutions are available for copy number analysis of FFPE samples. Both CNAG 2.0 and CNAT 4.0 correct for the fragment size bias in copy number predictions seen in FFPE data.

CNAG 2.0 was developed with specific features to accommodate FFPE samples and has been tested on a wide assortment of FFPE data (for example, see Figure 4a). This program can be downloaded for free from the following website: http://www.genome.umin.jp/

CNAT 4.0 also appears to work well with FFPE samples, although it currently has only been tested on a smaller sample size (for example, see Figure 4b). CNAT 4.0 can be downloaded for free from the following website: http://www.affymetrix.com/products/software/specific/cnat.affx

Exclusion of larger fragment sizes. The SNPs on larger fragment sizes are non-informative during copy number analysis of degraded samples and should be excluded. The option to filter out SNPs based on fragment size is included in both CNAG 2.0 and CNAT 4.0.

Reference set selection. As with all DNA samples, copy number analysis of FFPE DNA using references (normal, unaffected samples) that were processed in the same lab at the same time by the same user will provide the highest quality of copy number prediction. Using paired normal samples from the same patient, in the case of tumor DNA, is also often ideal for detecting somatic mutations associated with cancer.

Since paired and matched samples are not always available, both CNAG 2.0 and CNAT 4.0 allow the user to choose a paired or unpaired reference analysis. Either option can be used to create copy number profiles from FFPE samples (Figure 4), although reference choice may affect the level of noise seen in the results. When possible, normal references that were run in the same batch as the case sample are recommended. When no other options are available, it is possible to use the 48 HapMap samples available for download from the Affymetrix website: http://www.affymetrix.com/support/technical/sample_data/500k_data.affx

Conclusion/Summary
DNA obtained from FFPE samples can be applied to Affymetrix Mapping 500K Arrays for an integrated analysis of genotype, LOH and copy number. DNA quality varies greatly between samples, so a PCR-based QC step is strongly recommended to pre-qualify any FFPE DNA sample for the mapping assay. Below are a few additional recommendations to ensure successful processing and analysis of samples from FFPE blocks:

- Use Mapping 500K Arrays; do not apply FFPE DNA to Mapping 100K Arrays.
- When necessary, pool additional PCR reactions during the mapping assay to accumulate adequate amplicon products before proceeding to the fragmentation step.
• Filter SNPs during both LOH and copy number analysis so that larger fragments that were not amplified during PCR are excluded in downstream analysis.

• Only CNAG 2.0 or CNAT 4.0 should be used for copy number analysis of FFPE samples, but the same flexibility in reference choice is possible for these samples as are possible when using fresh DNA.

Acknowledgements

Much of this work is the result of collaborations with partners at the Peter MacCallum Cancer Centre (Melbourne, Australia), University of Tokyo (Tokyo, Japan), MD Anderson Cancer Centre (Houston, Texas), and University of Arizona (Tucson, Arizona). We wish to thank the following researchers for their contributions to this work:

Peter MacCallum Cancer Center, Melbourne, Australia:
  Ian Campbell
  Ella Thompson

University of Tokyo, Tokyo, Japan:
  Seishi Ogawa
  Yasuhiro Nannya
  Go Yamamoto

MD Anderson Cancer Centre, Houston, Texas:
  Gordon Mills
  Ralf Krahe
  Melissa Bondy
  Rahul Mitra
  Tuna Musaffe
  Aysegul Sahin
  Katherine Hale

University of Arizona, Tucson, Arizona:
  Patricia Thompson

Reference


*http://www1.qiagen.com/Products/GenomicDnaStabilization/Purification/DNeasyTissueSystem/DNeasyBloodTissueKit.asp
Notes: