MammaPrint Molecular Diagnostics on Formalin-Fixed, Paraffin-Embedded Tissue

This is the author's manuscript

Original Citation:
MammaPrint Molecular Diagnostics on Formalin-Fixed, Paraffin-Embedded Tissue / Sapino A; Roepman P; Linn SC; Snel MH; Delahaye LJ; van den Akker J; Glas AM; Simon IM; Barth N; de Snoo FA; van 't Veer LJ; Molinaro L; Berns EM; Wesseling J; Riley LB; Anderson D; Nguyen B; Cox CE. - In: THE JOURNAL OF MOLECULAR DIAGNOSTICS. - ISSN 1525-1578. - ELETTRONICO. - 16:2(2013), pp. 190-197.

Availability:
This version is available http://hdl.handle.net/2318/140944 since

Published version:
DOI:10.1016/j.jmoldx.2013.10.008

Terms of use:
Open Access
Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.
This Accepted Author Manuscript (AAM) is copyrighted and published by Elsevier. It is posted here by agreement between Elsevier and the University of Turin. Changes resulting from the publishing process - such as editing, corrections, structural formatting, and other quality control mechanisms - may not be reflected in this version of the text. The definitive version of the text was subsequently published in [The Journal of Molecular Diagnostics, 2013 December 28, DOI 10.1016/j.jmoldx.2013.10.008].

You may download, copy and otherwise use the AAM for non-commercial purposes provided that your license is limited by the following restrictions:

(1) You may use this AAM for non-commercial purposes only under the terms of the CC-BY-NC-ND license.

(2) The integrity of the work and identification of the author, copyright owner, and publisher must be preserved in any copy.

(3) You must attribute this AAM in the following format: Creative Commons BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en), [http://dx.doi.org/10.1016/j.jmoldx.2013.10.008]
MammaPrint Molecular Diagnostics on Formalin-Fixed, Paraffin-Embedded Tissue

Anna Sapino1, Paul Roepman2, Sabine C. Linn3, Mireille H.J. Snel2, Leonie J.M.J. Delahaye4¶, Jeroen van den Akker5, Annuska M. Glas6, Iris M. Simon7, Neil Barth8, Femke A. de Snoo9, Laura J. van ’t Veer2, Luca Molinaro1, Els M.J.J. Berns8, Jelle Wesseling9, Lee B. Riley10, David Anderson11, Bichlien Nguyen12, Charles E. Cox3

1 Department of Medical Sciences, University of Turin, Turin, Italy. email: anna.sapino@unito.it.
2 Department of Research and Development, Agendia, Amsterdam, the Netherlands.
3 Department of Medical Oncology, Netherlands Cancer Institute, Amsterdam, the Netherlands; Department of Pathology, University Medical Center Utrecht, Utrecht, the Netherlands.
4 Department of Product Development and Support, Agendia, Amsterdam, the Netherlands.
5 Department of Medical Affairs, Agendia, Irvine, California.
6 Department of Medical Affairs, Agendia, Amsterdam, the Netherlands.
7 Department of Medical Sciences, University of Turin, Turin, Italy.
8 Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, the Netherlands.
9 Department of Medical Oncology, Netherlands Cancer Institute, Amsterdam, the Netherlands.
10 Department of Surgical Oncology, St. Luke’s Hospital and Health Network, Bethlehem, Pennsylvania.
11 Department of Pathology, St. Luke’s Hospital and Health Network, Bethlehem, Pennsylvania.
12 Department of Medicine, Todd Cancer Institute, Long Beach Memorial Medical Center, Long Beach, California.
13 Department of Surgery, University of South Florida, Tampa, Florida.
MammaPrint, a prognostic 70-gene profile for early-stage breast cancer, has been available for fresh tissue. Improvements in RNA processing have enabled microarray diagnostics for formalin-fixed, paraffin-embedded (FFPE) tissue. Here, we describe method optimization, validation, and performance of MammaPrint using analyte from FFPE tissue. Laboratory procedures for enabling the assay to be run on FFPE tissue were determined using 157 samples, and the assay was established using 125 matched FFPE and fresh tissues. Validation of MammaPrint-FFPE, compared with MammaPrint-fresh, was performed on an independent series of matched tissue from five hospitals (n = 211). Reproducibility, repeatability, and precision of the FFPE assay (n = 87) was established for duplicate analysis of the same tumor, interlaboratory performance, 20-day repeat experiments, and repeated analyses over 12 months. FFPE sample processing had a success rate of 97%. The MammaPrint assay using FFPE analyte demonstrated an overall equivalence of 91.5% (95% confidence interval, 86.9% to 94.5%) between the 211 independent matched FFPE and fresh tumor samples. Precision was 97.3%, and repeatability was 97.8%, with highly reproducible results between replicate samples of the same tumor and between two laboratories (concordance, 96%). Thus, with 580 tumor samples, MammaPrint was successfully translated to FFPE tissue. The assay has high precision and reproducibility, and FFPE results are substantially equivalent to results derived from fresh tissue.

MammaPrint is a microarray-based in vitro diagnostic multivariate index assay for determining prognosis in early-stage breast cancer that uses the expression levels of 70 genes to assess the risk of recurrence in early-stage breast cancer. MammaPrint has been validated in an independent cohort of 302 patients from five European centers and in many additional validation cohorts, including lymph node–positive patients, HER2-positive patients, and patients with small tumors. The MammaPrint assay has been used to predict response to chemotherapy in adjuvant and neoadjuvant settings. Five-year outcome data in 427 patients, recently reported from a prospective observational study, demonstrated that 20% more patients are classified as low risk by MammaPrint, compared with standard clinicopathological stratification, without jeopardizing the outcome of those patients.

The MammaPrint test performed on fresh tissue (MammaPrint-Fresh) has been validated in a variety of studies involving more than 12,000 patients. Today, MammaPrint is routinely used in clinical practice around the world. The test is based on accurate measurement of gene expression by microarray analysis and has been developed and validated for fresh and fresh-frozen tumor tissue to guarantee high-quality RNA for the analysis. Gene expression analysis using FFPE tissue (a different state of the analyte) has long been challenging, because of nonstandardized fixation protocols and because of unreliable retrieval of high-quality RNA from FFPE material due to fragmentation of RNA and cross-linking as a result of formalin fixation.
In recent years, several new technologies have been developed for improving RNA extraction from FFPE tissue, making RNA amplification more efficient. Moreover, a joint American Society of Clinical Oncology and College of American Pathologists guideline issued in 2010 regarding the standards surrounding tissue handling in the preanalytic setting has significantly reduced unwanted variations in ischemia and fixation times, thereby facilitating the consistency of properly preserved tissue in the FFPE analyte state. An exploratory study in a subgroup of the MammaPrint genes indicated that the expression levels of these genes were preserved in FFPE samples if the improved techniques for RNA processing were used. Results from matched FFPE and fresh-frozen samples from the same tumor had a high correlation, suggesting that the analyte state does not substantially affect analytical and clinical performance.

Here, we describe method optimization and validation of the MammaPrint assay for use in FFPE tissue samples. Analytic performance was evaluated by assessing precision, repeatability, and reproducibility.

Materials and Methods

**FFPE Samples**

Technical optimization (RNA isolation amplification and cyanine dye labeling) of the assay was performed using a set of 157 FFPE samples as the technical cohort (Table 1). An additional set of 125 FFPE samples with matching fresh tissue samples was used as the establishment cohort, for calibration of the MammaPrint-FFPE read-out (Table 1). A third cohort, an independent equivalence cohort of 211 FFPE samples with matching fresh tissue samples, was used to validate the FFPE analyte test (Table 1). Clinical characteristics (institute, age at diagnosis, lymph node status, T-stage, estrogen receptor status, and HER2/neu status) of the equivalence cohort are presented in Supplemental Table S1. In the reproducibility cohort, precision, repeatability, and reproducibility were assessed by multiple processing of 87 FFPE tumors samples (Table 1). FFPE samples used in the present study had been fixed and stored in formalin for ≤2 years before the microarray analysis. All paraffin-embedded samples were sent to Agendia laboratories as tissue blocks or as 5- or 10-μm sections on coated glass slides. All FFPE samples were scored for tumor cell percentage by standard histological assessment. Samples that contained <30% tumor cells were (if possible) enriched by manual microdissection during the deparaffinization step of the RNA extraction procedure.
Table 1  Sample Cohorts Used for MammaPrint-FFPE

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Technical Establishment</th>
<th>Validation</th>
<th>Reproducibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohort size*</td>
<td>n Z 157</td>
<td>n Z 125</td>
<td>n Z 211</td>
</tr>
<tr>
<td>Tissue type</td>
<td>FFPE</td>
<td>FFPE and fresh</td>
<td>FFPE and fresh</td>
</tr>
<tr>
<td>Purpose</td>
<td>Technical optimization</td>
<td>Establishment</td>
<td>Validation</td>
</tr>
</tbody>
</table>

*Total Z 580.

Analyte Processing and MammaPrint Gene Profile Read-Out

RNA extraction was performed using two sections of 10-μm thickness or four sections of 5-μm thickness. Deparaffinization and total RNA extraction was performed using an RNeasy FFPE kit (Qiagen, Valencia, CA) according to the manufacturer’s instructions. RNA yield was quantified using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA) and RNA quality was checked using a PCR test based on abundance and fragment length of the glucose-6-phosphate dehydrogenase gene (G6PD), as described previously.23 Extracted RNA was amplified using a TransPLEX C-WTA whole-transcriptome amplification kit (Rubicon Genomics, Ann Arbor, MI).

Amplified cDNA was labeled using the Genomic DNA Enzymatic Labeling Kit (Agilent Technologies, Santa Clara, CA) and hybridized onto Agenda’s diagnostic arrays (custom-designed and produced by Agilent Technologies specifically for Agenda), both according to the manufacturer’s instructions. The algorithm to determine the MammaPrint-FFPE indices was equivalent to the algorithm used for current MammaPrint-fresh diagnostics. A linear calibration was applied to the FFPE-derived indices to match the low risk/high risk threshold of the MammaPrint fresh test, using a set of 125 samples and comparing the FFPE-derived indices with results for matching fresh-tissue samples, after which the assay was locked. The FFPE test was validated using an independent equivalence cohort of 211 tumor samples, for which both fresh and FFPE samples were analyzed.

Assessment of Reproducibility, Repeatability, and Precision

All validation experiments to assess technical performance of the tests were designed according to guidelines of the US Food and Drug Administration (FDA) (www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071148.htm, issued March 13, 2007; last accessed September 5, 2013) and the Clinical and Laboratory Standards Institute (CLSI; previously NCCLS).24 All technical validations were performed in samples classified as high risk, low risk, or close to the threshold of MammaPrint. Acceptance criteria, which were determined before performing the
experiments in the laboratory, accounted for biases introduced by various parameters such as equipment or batches of reagents, as well as operators.

Repeatability and precision of the assay were assessed using a precision evaluation experiment according to CLSI document EP5-A2,24 by which four FFPE samples were analyzed in duplicate over 20 consecutive days. Reproducibility was further evaluated by three MammaPrint control samples (one low-risk and two high-risk control samples) that had been analyzed during a period of 12 months to assess nearly all potential sources of variation.

The reproducibility of the FFPE test between two isolations was assessed using a panel of 55 tumors for which two different sections were processed in parallel. Interlaboratory reproducibility was assessed on a panel of 25 tumor samples that were processed in both Agendia laboratories (one in Amsterdam, The Netherlands, and the other in Irvine, CA).

Statistical Analysis

Analyses were performed using R statistical software version 2.14.1 (http://www.r-project.org), Analyse-it statistical analysis add-in version 2.30 for Microsoft Excel 2007 (Analyse-it Software, Leeds, UK), and MedCalc software version 12.6.1 (MedCalc Software, Ostend, Belgium). Agreement (equivalence) analysis between MammaPrint results for FFPE and fresh tissue was performed on the categorical outcome level in terms of concordance and Cohen’s κ score, as well as negative percent agreement (NPA, the proportion of MammaPrint-fresh low-risk samples that are also classified as low risk using FFPE tissue) and positive percent agreement (PPA, the proportion of MammaPrint-fresh high-risk samples that are also classified as high risk using FFPE tissue). Reproducibility, repeatability, and precision of the diagnostic assays were measured in terms of the relative stability (RS), calculated as $RS = 100 - RSD$, where $RSD$ is relative SD (ie, the standard deviation of the measurements as a percentile of the total MammaPrint range).

Results

Technical Procedures for FFPE Microarray Diagnostics

The technical procedures needed for gene expression profiling in FFPE tissue were evaluated and optimized in 157 breast cancer FFPE samples. Total RNA extraction yields were significantly associated with tumor sample size [$r = 0.45$; 95% confidence interval (95% CI), 0.29 to 0.58] (Figure 1A). For very high-quality RNA, the RNA quality assay results in four PCR products (91, 123, 145, and 177 bp) (Figure 1B). The
presence of PCR products shows that the cross-linking that can occur during the fixation process is reversed and that the length and quality of RNA fragments are sufficient for amplification for microarray purposes.

The MammaPrint diagnostic test requires mRNA transcripts that have been isolated from a tumor sample that contains at least 30% invasive tumor cells. The developed FFPE procedure allows for enrichment of tumor cells by microdissection before RNA extraction (Figure 1C). The amount of FFPE tissue should be equivalent to 50 mm² with a 10-μm thickness. Reproducibility of the microdissection procedure was validated using six samples with a tumor percentage of 20% (based on the whole section) that were enriched to 60% to 70% tumor cells. The microdissections were performed in duplicate by two technicians. After isolation, amplification, and hybridization, the duplicate microdissections showed identical diagnostic outcomes (6/6) and almost identical indices ($r = 0.964$; 95% CI, 0.545 to 0.998) (data not shown).
Amplified cDNA was labeled with Cy3 dye and hybridized on Agendia’s diagnostic array. FFPE gene expression signals were checked for quality issues using a microarray quality-control model that is analogous to the quality model used for fresh diagnostics. Amplification, labeling, and microarray analysis of FFPE RNA resulted in a 97% success rate.

**MammaPrint Gene Profiles in FFPE Samples**

MammaPrint 70-gene profiles of 10 FFPE samples for which a matched fresh tissue sample indicated a low-risk result were compared with those of 10 FFPE samples with a high-risk result. As expected, gene expression levels differed significantly between low-risk and high-risk FFPE samples (Figure 2A). A direct comparison of the MammaPrint 70-gene profile between the FFPE and the matching fresh tissue samples indicated a very strong concordance, both for the low-risk ($r = 0.881; 95\%\ CI, 0.815\; to\; 0.925$) and the high-risk ($r = 0.832; 95\%\ CI, 0.743\; to\; 0.893$) profiles (Figure 2B).

![Figure 2](image-url)

**Figure 2**

MammaPrint gene profiles in FFPE. A: MammaPrint 70-gene heatmap profiles on FFPE samples for which the matching fresh tumor samples indicated a high-risk (n = 10) or low-risk (n = 10) outcome. B: Direct comparison of MammaPrint 70-gene expression between FFPE and fresh analyte-based analysis. For each of the 70 MammaPrint genes, the low-risk and high-risk aggregate expression across the 10 samples in A is compared against matching fresh-tissue expression levels (median centered). C: Stability of the control probes assessed on eight independent FFPE samples. Correlation in gene expression levels across the normalization probes is shown for four random pairs.
The normalization genes (or reference genes) had been selected to show a stable expression between samples, independent of MammaPrint outcome. The stability of the normalization control genes was assessed on eight FFPE samples, randomly grouped into four pairs; the assessment indicated a very high correlation in gene expression levels between pairs, with an average r-value of 0.996 (95% CI, 0.995 to 0.997) (Figure 2C).

**Equivalence of MammaPrint Results for FFPE and Fresh Tissue**

Equivalence of MammaPrint for FFPE and for fresh tissue was validated using a set of 211 samples from five hospitals with available matching fresh and FFPE tissue samples. Clinical characteristics of the patients [median age, 60 years (range, 28 to 93 years), 80% lymph node negative, 86% ER positive, 9% HER2/neu positive] (Supplemental Table S1) are representative of patients whose samples were used for development and validation of the MammaPrint-fresh assay. MammaPrint-FFPE indices showed a high correlation with the matching indices for fresh tissue samples ($r = 0.917$; 95% CI, 0.893 to 0.936) (Figure 3). Concordance of categorical low-risk and high-risk classification was 91.5% (95% CI, 86.9% to 94.5%) in this validation set; the NPA was 91.3% (95% CI, 84.4% to 95.4%) and the PPA was 91.6% (95% CI, 84.8% to 95.5%). A $\kappa$ score of 0.82 (95% CI, 0.74 to 0.89) indicated "an almost perfect agreement."25 The majority of discordant samples (14 of 18) were close to the MammaPrint threshold (within ±5% of the diagnostic range) for either the fresh or the FFPE index. For samples within these 5% ranges ($n = 48$), the NPA was 76.7% (95% CI, 66.0 to 85.4) and the PPA was 61.1% (95% CI, 43.3 to 75.6). For all other samples ($n = 163$), the NPA was 97.3% (95% CI, 93.2 to 98.9) and the PPA was 97.8% (95% CI, 94.4 to 99.1).

![Figure 3](Equivalence of MammaPrint results from matched FFPE and fresh tissue samples in an independent validation cohort. Equivalence of MammaPrint-FFPE was assessed on tumor samples for which both the FFPE and fresh tumor samples were analyzed ($n = 211$). FFPE-derived indices are plotted against indices from the matching fresh tissue samples. Low-risk versus high-risk categorical thresholds are indicated by dashed lines.)
Precision and Reproducibility of MammaPrint-FFPE

The reproducibility of the FFPE assay was assessed using multiple samples from 55 tumors, each tumor sectioned twice and the two sections (A and B) were processed in duplicate on different days. MammaPrint indices showed a very strong correlation between replicate analyses ($r = 0.972$; 95% CI, 0.953 to 0.984), with a concordance of 96% (Figure 4A). These results confirm high reproducibility on multiple independent analyses of the same tumor.

![Figure 4](image-url)
Discussion

Compared with RNA extracted from FFPE tissue, the quality of RNA extracted from fresh or fresh-frozen tissue is superior and thus is generally considered the most suitable for identification of robust biomarkers and gene profiles. In recent years, however, preanalytic standards and analytic technologies have greatly improved, allowing for the extraction and processing of possibly chemically modified (cross-linked) and partially degraded RNA and thereby allowing high-quality microarray assays on FFPE tissue.

In our laboratory, after identification of optimal methods for FFPE RNA extraction, amplification, and labeling, microarray hybridization from FFPE samples had a success rate of 97%. The MammaPrint test for the FFPE analyte resulted in an overall concordance of 91.5% (95% CI, 87% to 95%) on 211 independent matched FFPE and fresh tumor samples. The technical performance of the test thus has a high concordance with that reported for the MammaPrint-fresh assay (κ = 0.829; 95% CI, 0.754 to 0.905).

The concordance rate of 91.5% observed in the present study reflects intrinsic differences in matched tumor tissue samples, as well as assay variability. Quantification of tumor heterogeneity is challenging, because in a diagnostic setting typically only a single biopsy can be analyzed. For the MammaPrint-fresh assay, we have previously observed 95% equivalence when analyzing two biopsy isolations from the same tumor. These results indicate that a discordance of 5% between MammaPrint test results for the same tumor can likely be allocated to cellular heterogeneity within the primary tumor. The additional 3.5% discordance observed in the present study can be ascribed to the FFPE assay; that is, a substantial part of the discrepancy between fresh and FFPE samples is contributed by tumor heterogeneity rather than assay variability. Clinically, this discordance means that the chance of the test result (from a fresh block) being different when run a second time from a given FFPE tissue block is 8.5%. Our results indicate that most of the patients with a discordant result had a MammaPrint index that falls within ±5% of the diagnostic range of the classification threshold, with a similar distribution of low-risk and high-risk results. If the analytical accuracy of the test result is <90%, the physician is notified of a borderline result on the result form. The assay is intended to supplement information provided by clinicopathologic factors, and physicians should account for accuracy variance in their decision making.

Microarray technology as a platform for molecular diagnostics provides significant advantages over more conventional technologies such as reverse transcription-PCR and immunohistochemistry, largely because the number of genes that can be read out from one patient tumor sample is nearly unlimited. This allows parallel read-out of multiple markers and profiles using only a single biopsy.
In addition to signature genes, the microarray platform allows a large number of control and normalization genes to be assessed simultaneously under the same conditions. Thus, only one laboratory reaction (a single sample hybridization) establishes the quality of multiple tests and allows an almost unlimited number of genes to be interrogated.

Several other multigene assays are currently available for early breast cancer patients; one of these, PAM50, was originally developed on fresh-frozen tumor samples but later became commercially available for FFPE tumor samples.32 To our knowledge, however, no previous report compares the two test analytes, FFPE versus fresh, or describes assay adaptation for the different analytes.

Clinical utility (and physician adoption) of molecular tests can be optimized only if the tests are provided as robust, standardized, preferably centralized assays with documented high accuracy, repeatability, precision, and reproducibility in a number of different analyte species. If technical variation is to be minimized, then diligent assessment of all possible unwanted sources of variation and bias (such as scanners, reagents, and operators) is an important aspect of technical development of molecular assays with different analytes.

In conclusion, the MammaPrint test can be used on core and surgical sections from FFPE tissue as an alternative to fresh tissue. The MammaPrint-FFPE assay has excellent reproducibility, precision, and repeatability, with performance closely similar to that of MammaPrint-fresh.

Acknowledgments
We thank Arno Floore, Diederik Wehkamp, Jeroen Heijmans, and Frank Nieboer for their contributions to this study.

References


Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–174


