ELSEVIER

Cancer Genetics 205 (2012) 341–355

Cancer Genetics

REVIEW

Molecular inversion probes: a novel microarray technology and its application in cancer research

Yuker Wang [a], MariEllen Cottman [a], Joshua D. Schiffman [b,*]

[a] *Assay Development and Bioinformatics Department, Affymetrix, Inc., Santa Clara, CA, USA;* [b] *Center for Children's Cancer Research (C3R), Departments of Pediatrics and Oncological Sciences, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA*

The molecular inversion probe (MIP) assay technology was originally developed for single nucleotide polymorphism (SNP) genotyping, but has subsequently been used for identifying other types of genetic variation including focal insertions and deletions, larger copy number alterations (CNAs), loss of heterozygosity (LOH), and most recently, for somatic mutation detection. The assay requires as little as 75 ng of genomic DNA and has been shown to perform well with highly degraded DNA, such as that from formalin-fixed paraffin-embedded (FFPE)-preserved samples from 20 years ago or older. Central to the MIP assay technology are the padlock probes that hybridize to the DNA target of interest before polymerase chain reaction amplification, leading to high assay specificity. As outlined in this review, the MIP assay has enabled new discoveries and a deeper understanding of the molecular basis of cancer and its various disease subtypes. The use of novel genomic technology such as MIPs on clinically archived FFPE samples has the potential to lead to more accurate disease diagnosis, prognosis, and novel therapeutic intervention. This review describes the initial history of MIP technology, details of the MIP assay, its current analysis techniques, and recent publications related to this novel platform.

Keywords Molecular Inversion Probes (MIPs)' cancer genomics, microarray

Molecular inversion probe (MIP) technology is a relatively new and powerful single nucleotide polymorphism (SNP) microarray assay platform that continues to gain momentum in both research and clinical cancer genomics. The first applications of the MIP assay technology were for genotyping SNPs. Subsequently, the MIP assay has been used for identifying other types of genetic variation including focal insertions and deletions, larger copy number alterations, loss of heterozygosity (LOH), and most recently, for somatic mutation detection. This review describes the initial history of MIP technology, details of the MIP assay along with current analysis techniques, and recent publications related to this novel platform.

History and development of MIP technology

MIP assay technology has an exciting history based on cutting-edge technology and the determination of research scientists to continually improve modern-day genotyping assays. The MIP assay uses a nucleic acid probe that hybridizes to a genomic target, so that the two ends of the probe hybridize to the target in two locations spanning an ~40 basepair (bp) footprint. The ends of the probe can then be subsequently joined to create a closed circular probe, or padlock probe, which was first described in 1994 (1). The probes were later referred to as molecular inversion probes or MIPs. The closed circular probe is resistant to exonucleases and can be enriched relative to other nucleic acids in the mixture and then amplified using universal primers. Scientists at the Stanford Genome Technology Center in Palo Alto, California, further developed the method for cleavage of the probe followed by amplification, for example, by polymerase chain reaction (PCR) (2). The amplification products can then be detected by hybridization to an array of probes complementary to the padlock probes.

These innovative developments eventually led to the founding of a company called ParAllele Bioscience in 2000.

Affymetrix®, DMET™, and OncoScan™ are trademarks or registrered trademarks of Affymetrix, Inc. and the assays discussed are for research use only.

Received May 9, 2012; received in revised form June 9, 2012; accepted June 11, 2012.

* Corresponding author.

E-mail address: Joshua.Schiffman@hci.utah.edu

2210-7762/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cancergen.2012.06.005

Dr. Paul Hardenbol and a team of scientists at ParAllele Bioscience continued to develop and streamline the MIP assay technology (2). The unique homology of the tag sequences allowed for pooling of the probes, thus enabling high levels of assay multiplexing. The technology rapidly evolved and the SNP multiplex levels increased from 1,000-plex in 2003 to >10,000-plex in 2005.

Participation in the International HapMap project

ParAllele Bioscience was chosen to be a member of the International HapMap Consortium in 2002, and the MIP assay technology was used to genotype thousands of SNPs on chromosome 7 and then >11,000 non-synonymous SNPs as part of the HapMap project (3). As an easy-to-customize genotyping platform, the MIP assay technology had been used by many investigators as a targeted genotyping platform for custom applications ranging in number from 3,000 to 10,000 SNP markers. One of the first available SNP genotyping panels was a human 20,000 non-synonymous coding SNP panel. As part of a collaboration with Eli Lilly and Company (Indianapolis, IN), ParAllele Bioscience designed and created a 5,000-probe panel that targeted drug metabolism genes relevant to pharmacology (4,5). Other MIP assay panels for SNPs within inflammatory genes were also available, but the focus for MIP development became the cancer panels discussed in the 'MIP assay for cancer application' section.

Acquisition of ParAllele Bioscience by Affymetrix, Inc.

In October 2005, Affymetrix, Inc. (Santa Clara, CA) acquired ParAllele Bioscience and continued to develop the technology with two focus areas for application development: pharmacogenetics and cancer copy number analysis. Over time, the MIP assay was found to perform quite well in formalin-fixed paraffin-embedded (FFPE) archived samples, which is one of the advantages of this particular SNP array technology for cancer investigation.

In 2008, Affymetrix launched the DMET Plus Solution, which uses the MIP assay technology to detect 1,936 complex pharmacogenetic variants; the DMET Plus Assay is a comprehensive single, multiplexed assay for accurately and reliably analyzing known markers for drug metabolism studies. This DMET Assay was recently demonstrated to work quite well in DNA extracted from saliva (6), further enabling the study of pharmacogenetic-related toxicities.

In addition to pharmacogenetic applications, Affymetrix focused significant effort on the development of MIP technology for copy number, LOH, and somatic mutation analysis in cancer research. Initially, the MIP assay contained a 24,037 probe cancer panel which focused on cancer genes that included tumor suppressors, oncogenes, and DNA repair genes. This initial 24K cancer panel also included coverage across the entire genome, resulting in median probe spacing every ~43 kilobases (kb). After substantial beta-testing, the MIP assay was introduced by Affymetrix in September 2010 as the "OncoScan FFPE Express Service." This newly expanded assay now contained 330K probes, with a focus still on cancer genes. The additional probes provided even denser coverage within cancer genes of interest and still offered good coverage across a genome-wide backbone with a total median probe spacing of 4.2 kb. The most recent product, the OncoScan FFPE Express 2.0 Service, was launched in April 2011. This latest iteration of the MIP assay offered further enhancement of coverage for tumor suppressors and oncogenes. The new OncoScan FFPE Express 2.0 was designed in collaboration with members of the cancer research teams formed by the American Association of Cancer Research (AACR) Stand Up to Cancer (SU2C) initiative, with further input from international clinical and basic scientists in the field of cancer research. The current assay interrogates 335K markers for copy number changes, LOH, and 541 somatic mutations specific for cancer. Specifically, the OncoScan FFPE Express 2.0 offers dense coverage of over 200 tumor suppressor and oncogenes, with a median spacing of 1 probe per 0.5 kb for the top 10 "actionable" tumor suppressor genes, a median spacing of 1 probe per 2 kb for the top 190-plus actionable oncogenes, and a median backbone spacing of 1 probe per 9 kb. A large amount of effort went into designing this latest version so that it would include genes and mutations in the PI3K pathway, which is the focus of one of the SU2C teams.

Cancer Cytogenetics Microarray Consortium (CCMC) study

The CCMC is comprised of clinical pathologists, cytogeneticists, and molecular genomicists whose mission is to promote communication and collaboration among cancer cytogenetics laboratories, and their 2011 Annual Meeting has been reviewed in this journal (7). Specifically, the CCMC has undertaken steps to 1) set up platform-neutral standards for cancer microarray designs and create cancer-specific platforms suitable for cancer diagnosis; 2) share cancer microarray data; 3) create a public cancer array database; and 4) conduct multicenter cancer genome research. As part of the CCMC's initial validation studies of different microarray technologies, the MIP platform was chosen as one of the candidate platforms. Thus far, the CCMC has used MIPs to investigate FFPE and matched fresh-frozen renal samples. Although the final results have not yet been analyzed, the initial findings from the CCMC's use of MIP assay technology is quite promising with very high-quality results comparable to other standard microarray platforms currently in use (Personal communication, Brynn Levy).

MIP assay design

The version of the padlock probe assay described by Hardenbol et al. (2) had several unique features that ensured increased specificity and accuracy of genotyping: 1) four independent nucleotide specific reactions for each nucleotide base (A, C, G, T); 2) use of the sequence-specific Stoffel fragment of polymerase, which lacks the 5′-3′ exonuclease activity; 3) sequence-specific thermostable DNA ligase enzyme; and 4) cocktail of DNases to digest unused probes and genomic DNA while leaving the circularized padlock probe intact. See Figure 1 for details of the MIP assay before and after the inversion steps that utilize these unique designs.

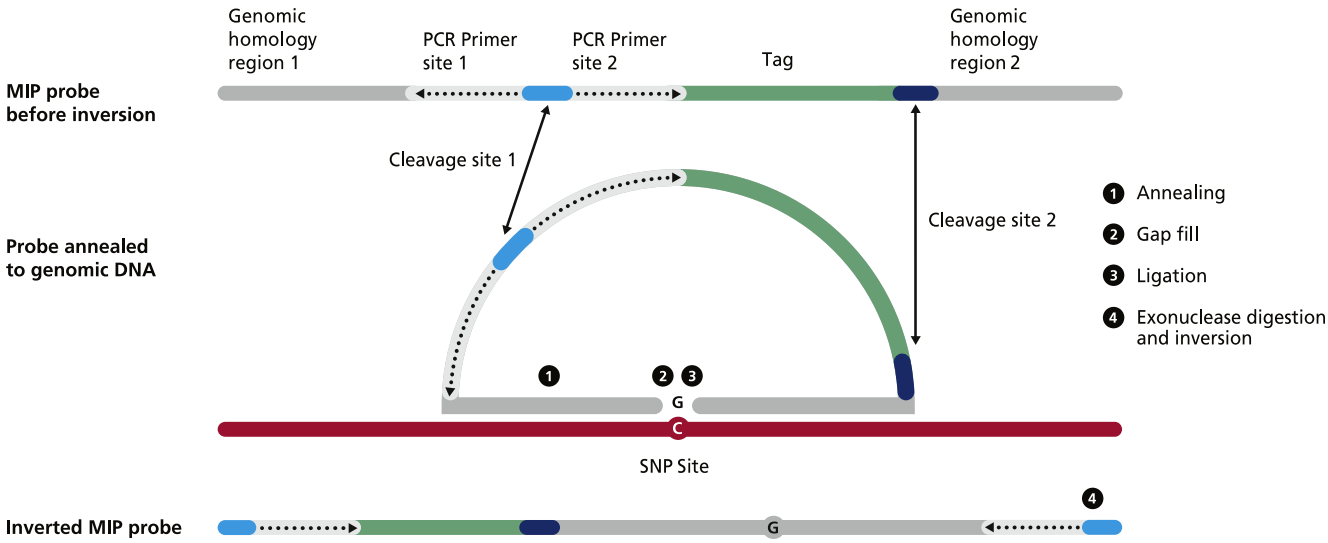

Figure 1 The molecular inversion probe before, during, and after inversion. (Top) The pre-inverted MIP probe including: genomic homology regions, universal PCR primer binding sites, unique Tag, and cleavage sites. (Middle) The MIP probe annealed onto a genomic DNA template after annealing and gap-fill at the SNP site (in this example, the complementary base is a T). The circularized probe is released from the DNA, and then all single stranded material is destroyed by exonuclease. The probe is then digested at cleavage site #1 and it becomes inverted. (Bottom) The inverted probe is amplified with universal primers.

From the beginning, the MIP assay technology was considered a robust genotyping platform that generated very high-quality data with low noise and background compared to other high-throughput multiplex genotyping platforms (2). This increase in genotyping performance can be largely attributed to the dramatic reduction in genome complexity before hybridization to the array (for example, step 4, exonuclease digest in Figure 1). In addition, the MIP assay does not require whole genome amplification of the initial genomic DNA, a process that can introduce substantial noise into genomic array analyses (8). Complexity reduction translates to a more specific signal, less noise from cross-hybridization, and better signal-to-noise ratio resulting in a greater dynamic range after the signal is transformed to copy number.

Another advantage of the MIP assay is its ability to provide high-quality, high-resolution results for both genotype and copy number from as little as 75 nanograms (ng) of genomic DNA. In fact, many investigators have reported excellent quality results from the MIP assay from as little as 20 ng of DNA, or in some cases, even less.

Detection for both SNPs and copy number aberrations

As described, the MIP assay technology was originally used for SNP genotyping, but with time it became clear that it could be used for more quantitative assessments such as copy number quantification. In the first publication to describe the quantitative aspects of the MIP assay, the “allele ratio” was introduced, which is a method to identify heterozygous SNPs within copy number regions as well as a way to define the absolute copy number (9). This same publication also described the use of allele information to show mosaic abnormalities over large genomic segments. This analysis technique demonstrated a large-scale, copy-neutral mosaic of 97 megabases (Mb) on chromosome 15, as illustrated in figure 5 of reference (9). In order to understand the copy number performance quantitatively, and the trade-off between resolution and sensitivity, another analysis method was introduced. A receiver-operating characteristic (ROC) curve was used to assess the change between CN $=$ 1 and CN $=$ 2 (i.e., male chrX and autosomes) (9). ROC analysis is still considered the best method to assess copy number performance (10). This initial MIP copy number publication from 2005 also included a proof of principle experiment for using FFPE samples with the MIP assay. The FFPE samples included brain, colon, and liver tissue, and were between 1 to 3 years of age. These FFPE samples contained varying levels of DNA degradation, but the experimental data revealed surprisingly high-quality results. As explained in the original publication (9), the high-quality results from FFPE are due to the fact that the individual probes in the MIP assay bind to a genomic footprint of roughly 40 base pairs (bp); therefore, degraded DNA is less of an issue than in other SNP genotyping assays that often require the amplification of regions several hundred bps to $>$1 kb in size. The performance in degraded FFPE samples has been a unique feature of the technology since then and is one of the advantages of the MIP assay over traditional SNP genotyping assays.

Two years later, another technical assessment using MIP assay for copy number variation was published (11). In this paper, it was described that an impressive false-positive rate at 10^{-4} could be achieved by smoothing two markers in a 50K MIP assay panel. Furthermore, copy number assessment was compared with a series of spike-in

materials. In this experiment, linear and unsaturated copy number calls of up to 60 copies were achieved with very little cross-hybridization between the two alleles (e.g., allele A remained CN = 1, while allele B reached CN = 120). Also demonstrated was the ability of MIP assays to identify LOH without matched normal tissue and the remarkable ability for the assay to work reliably on limited amounts of initial, start-up genomic DNA down to 75 ng.

After another 2 years, a third technical paper was published that highlighted the cancer research application of MIP assay with FFPE samples, along with a new data quality metric for copy number calls (12). A ROC curve is useful as a copy number quality control metric, but has limitations when applied to tumor samples with unknown aberrations. Wang et al (12). have described a new analysis approach, 2p-RSE, for use with the MIP assay data. 2p-RSE takes the median of the relative standard error of a whole genome pairwise comparison and uses the median value to avoid counting abnormal breakpoints that are frequently detected in tumor samples. This method can apply to all samples—not just those with previously known copy number changes. This new approach is very useful when measuring the quality of genomic results from FFPE samples with varying levels of DNA degradation. The publication also demonstrated the continued utility of the allele ratio to confirm copy number changes and copy-neutral LOH in tumor samples. Also highlighted was the ability of the allele ratio to infer normal stromal tissue contamination, a problem that we now recognize to occur in almost all tumor samples. See Figure 2 for an example of inferring copy number calls from the allele ratio in the presence of normal tissue, and also see figure 3 in Wang et al. (12) for detailed discussion on this subject.

As part of this same recent MIP assay technology paper, a wide range of clinically archived FFPE samples (93 tumors from 7 collections) were collected and tested on a 50K-probe assay (12). These samples comprised several different tumors, including bladder, breast, colorectal, kidney, and liver cancer, and these FFPE blocks ranged in age from 5 months to over 20 years. The collection methods varied for these samples, as well as the processing methods (in fact, one cohort of samples were hematoxylin-eosin stained with subsequent microdissection). Using only 37 ng of genomic DNA, the FFPE samples produced an overall 88% sample pass rate. The genotype results were compared between matched fresh-frozen tissue

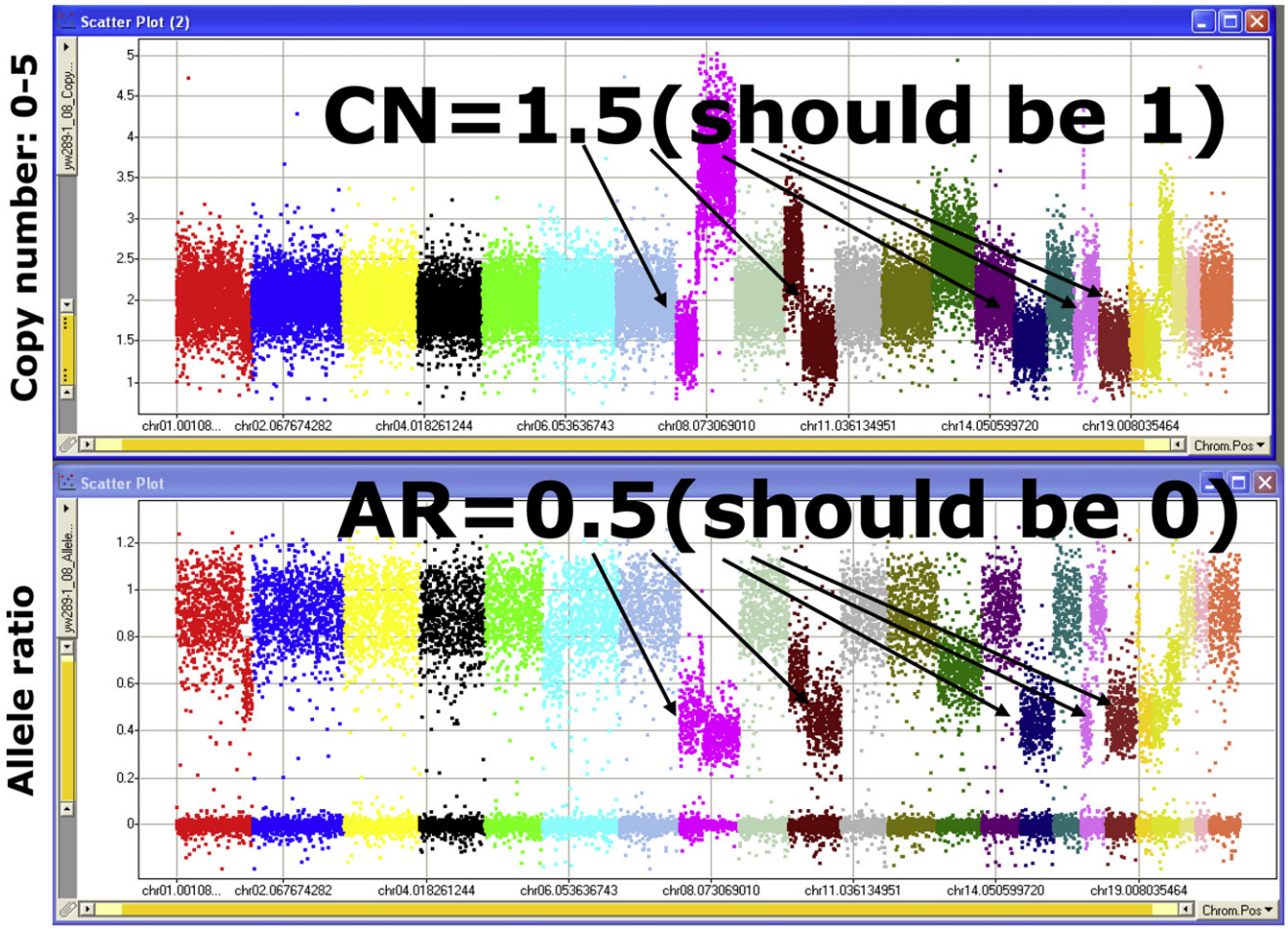

Figure 2 Copy number (CN) and allele ratio (AR) displayed from the same tumor in the top and bottom panels, respectively. Markers are arranged along the chromosomes and colored by chromosome. There is no smoothing or segmentation. In this sample there is ~50% normal (therefore CN = 2 in the arrowed region) and ~50% tumor. In the arrowed regions, the tumor appears to have one copy only. After mixing, the CN call becomes 1.5. The AR confirms this hypothesis. The heterozygous cluster for normal two copy should be 1 (1A/1B), whereas for tumor one copy, there is none (0A/1B). After mixing, the AR becomes 0.5 (1A + 0A/1B + 1B).

and FFPE samples with an average concordance rate of 99.9%, demonstrating the robustness of the MIP assay on very old and degraded samples.

MIP assay for cancer application

Working with cancer samples presents several important challenges, some of which can be addressed by the MIP assay and its design: 1) Sample heterogeneity: Clonal expansion plus genomic instability can result in a highly heterogeneous cellular population within tumors that can be difficult to identify. This issue can be addressed by looking at the level of the single cell or by digital counting, but both are time-consuming and costly. The MIP assay can help to address these issues of clonal heterogeneity by enabling a wide copy number dynamic range along with the allele ratio information generated by SNP genotyping. 2) Limited quantity of starting DNA from clinical samples: The MIP assay allows DNA inputs as low as 75 ng on a routine basis, which makes the technology ideal for use with limited clinical research samples, such as core needle biopsies and/or very limited and precious samples. 3) FFPE samples: Highly degraded FFPE samples can provide very important information for both basic and clinical scientists, but older specimens have been very difficult to interrogate with modern microarrays or next generation sequencing (NGS). It is estimated about 1 billion FFPE blocks are banked around the world (13). Using medical records in association with the long life span of the blocks enables highly powered, retrospective studies that can accelerate new drug development if genomic information can be coupled with retrospective drug response, such as in companion diagnostic tests. These FFPE blocks produce poor-quality results with many of the currently available SNP genotyping assays or NGS technologies. As outlined previously and discussed in the section "Publications using MIP assay technology on clinical samples," MIP assay technology can interrogate genomic DNA extracted from archived FFPE specimens at least 20 years old with reliable high-quality results.

Details of the MIP assay and related bioinformatics

The MIP assay is currently available as a cancer panel and offered as a research service from the Affymetrix Research Services Laboratory as the OncoScan FFPE Express 2.0 Service (Affymetrix, Inc, Santa Clara, CA). An off-the-shelf product is in development for clinical users and as described previously, the CCMC has begun work to validate the assay for use in clinical cytogenetic laboratories. Over 6,000 samples have been processed to date at Affymetrix from over 100 tissue types, and the sample pass rate continues to exceed 90%. Thus far, the assay has been successful in samples as old as 28 years, and older samples are now being acquired for interrogation by the MIP assay. As detailed below, Affymetrix has partnered with BioDiscovery, Inc. (El Segundo, CA) to offer software programs, such as Nexus Copy Number software, which enables the rapid analysis and correlation of copy number and somatic mutations data with mRNA, miRNA, and clinical outcome data. One advantage of the Nexus Copy Number software is that it has been designed to analyze MIP assay data and does not require customers to invest in significant bioinformatics resources for this unique microarray platform.

As outlined in the previous section, MIP assay technology is unique in its ability to accurately discriminate between copy number states from 0–60 copies, and sometimes even higher (11). The MIP assay enables a high signal-to-noise ratio for optimal discrimination power from degraded samples. Figure 3 shows brain tumor data illustrating an over 60-fold amplification in copy number for *PDGFRA*, which has been validated by corresponding gene expression of similar value (J.D.S., unpublished).

Another advantage of the MIP assay is the ability to detect copy-neutral LOH in degraded FFPE samples. This genetic aberration can only be detected with an SNP-based assay technology, and as shown in Figure 4, allele ratio data easily reveals copy-neutral LOH. One of the advantages of SNP microarrays compared to the older array comparative genomic hybridization (aCGH) technologies is this ability to detect allele-specific copy number and allelic imbalances for individual SNPs. It has been demonstrated that as much as 50–75% of the LOH observed in cancer samples is copy-number neutral versus a loss in genomic material (14–16). The incorporation of copy-neutral LOH data into future studies with clinical correlations continues to gain importance in cancer research.

The MIP assay (OncoScan FFPE Express 2.0 Service)

The process of running the MIP assay is fairly straightforward and begins with extraction of genomic DNA from the samples of interest. Oftentimes, a commercially available kit yields sufficient quality and quantity of DNA for this platform when using either fresh-frozen or FFPE specimens. At least 75 ng of genomic DNA is recommended per sample as measured by PicoGreen (Life Technologies, Carlsbad, CA) analysis (which gives more accurate results than optical density [OD] quantification). Although 75 ng is the recommended input amount, reliably high-quality results have been achieved in samples with as low as 20 ng of DNA and, in some situations, as low as 3 ng of DNA from FFPE blocks (J.D.S., unpublished). The researcher aliquots the genomic DNA samples into empty barcoded PCR plates provided by Affymetrix. The plates are then shipped to Affymetrix where they are tested for DNA quantity and quality, and a QC report is then sent back to the researcher. Two quality control steps are built into the assay process to check whether enough DNA is available for successful detection of copy number and somatic mutations post–array hybridization. The first quality control step is PicoGreen analysis performed as soon as the samples are received and before starting the assay. The second quality control step involves an agarose gel to determine sample quality after the first PCR stage of the assay. The absence of the appropriate band on the gel tends to correspond with poorer sample performance. The researcher may choose at this point whether or not to continue with the processing of samples lacking the appropriate bands, as they occasionally perform well in the assay. Once past these quality control steps, the DNA samples are

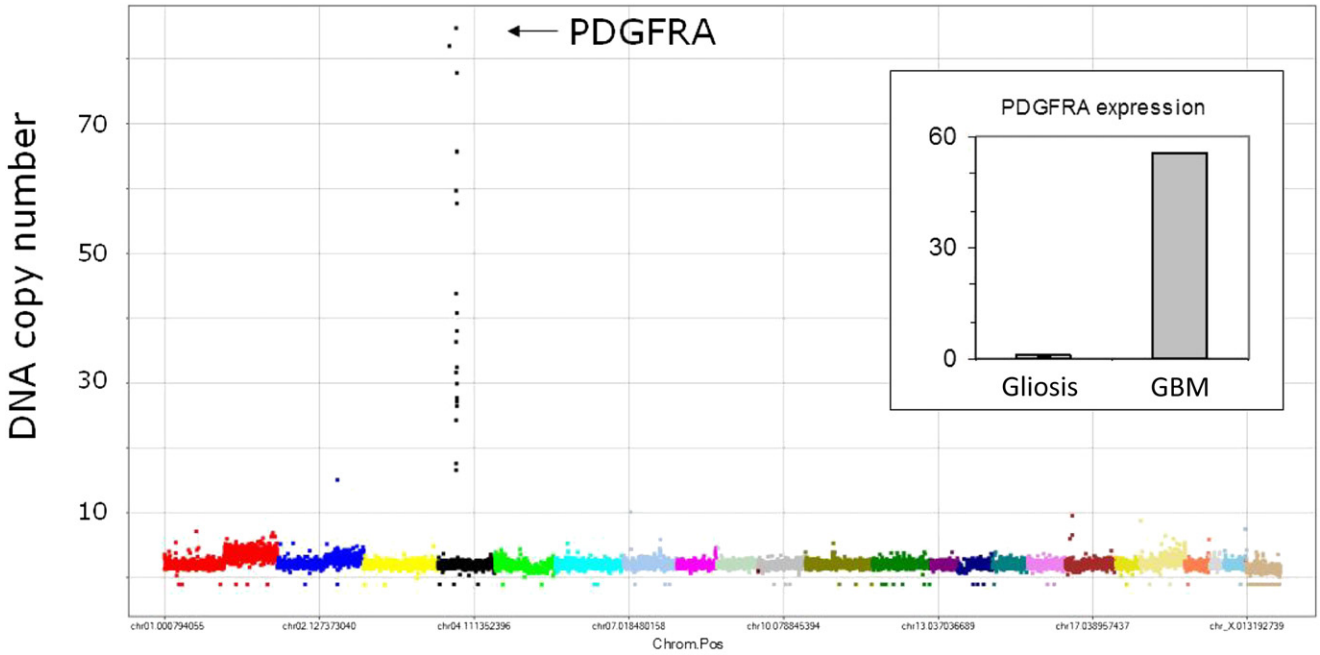

Figure 3 Data from pediatric glioblastoma (GBM) sample run on MIP assay illustrating a broad dynamic range for copy number changes. This plot shows a very high copy amplification of nearly 80 copies (linear) of the *PDGFRA* gene. The markers of this MIP 24K assay cancer panel are represented as individual data points, with each chromosome highlighted in a different color. No smoothing or segmentation is applied. Of note, *PDGFRA* gene expression in this GBM sample corresponds to the high copy number gain when compared to the same gene expression in non-cancer brain tissue (gliosis).

further processed using the MIP assay and then hybridized to the MIP array. Resulting data are reviewed by a bio-informatician at Affymetrix and delivered to the researcher along with the OncoScan FFPE Express 2.0 Service data analysis tools, including files formatted for Nexus Copy Number software analysis.

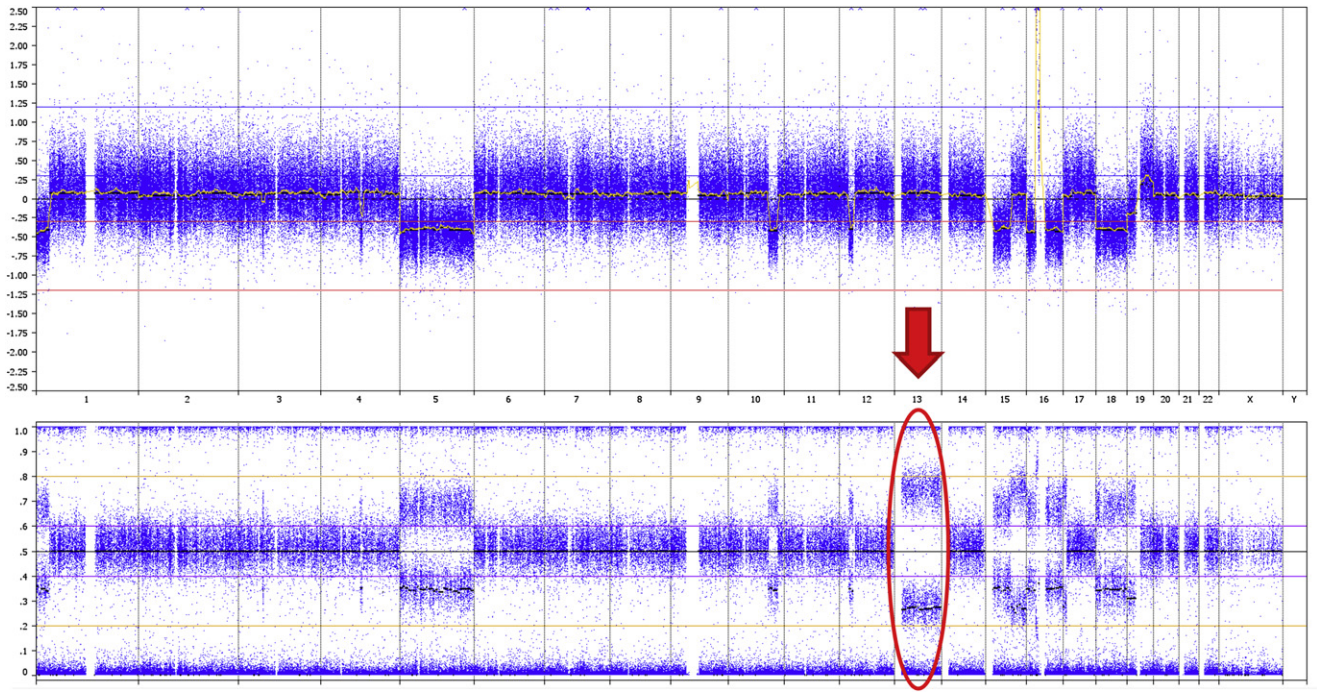

Figure 4 This figure illustrates copy-neutral LOH on chromosome 12 as displayed in Nexus Copy Numbersoftware version 5.1 (BioDiscovery, Inc., El Segundo, CA). (Top) Copy number data. (Bottom) Allelic ratio data (a summary of SNP information, which provides the ability to measure copy-neutral LOH). The allelic ratio information shows that both copies of chromosome 12q are identical (homozygous), leading to an increased likelihood of mutation susceptibility, as well as a possible locus of oncogene activation or tumor suppressor inactivation.

The MIP assay is a multiplex technology, and hundreds of thousands of probes are pooled into each initial reaction. Figure 5 illustrates the first stage and subsequent biochemistry of the MIP assay. Each probe in the MIP assay targets a unique SNP or base pair of interest. There are four initial reactions, one for each nucleotide. After this gap-fill step, there are several further reactions for exonuclease selection, probe release, and amplification. Once these previously bound probes are amplified, they are labeled with fluorescent molecules to enable detection. The "A" and "T" reactions are combined and the "G" and "C" reactions are combined; all reactions are then hybridized to the MIP arrays.

Two other reviews have been published, which also describe the technical details of MIP assay technology and how samples are processed (17,18).

Copy number data analysis

The signal from the array is normalized, transformed, and a clustering algorithm is applied on a marker-by-marker basis.

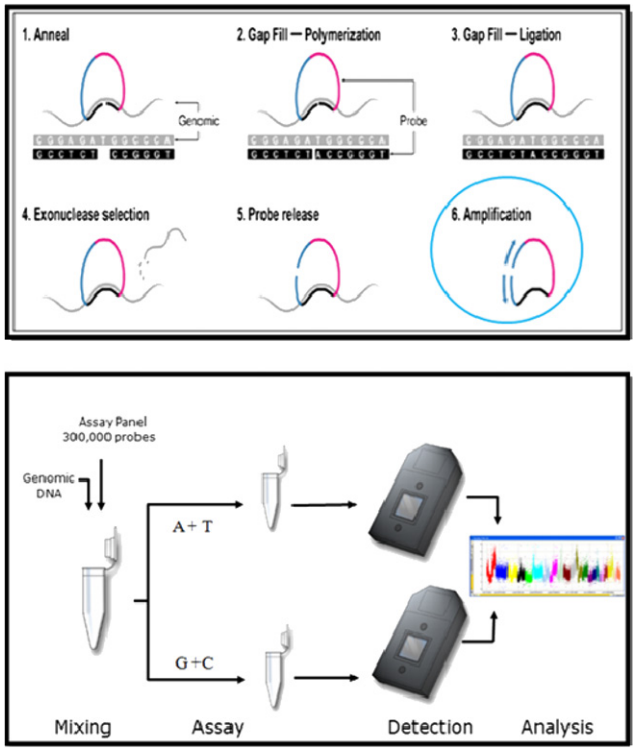

Figure 5 The biochemistry and workflow for the MIP assay is displayed. (Top) (1) The molecular inversion probe anneals to the genomic DNA. (2) A complementary nucleotide fills the gap at the site of the SNP. (3) The nucleotide is enzymatically ligated to the probe. (4) Exonuclease destroys single stranded material in the reaction. (5) The probe is cleaved and becomes inverted. (6) The probe is amplified using universal primers. Note that the captured probes are amplified during this step and not the original genomic DNA. This is one of the main reasons why the MIP assay works well with degraded FFPE DNA. (Bottom) A high-level overview of MIP technology processing. Mixing takes place on day 1, assay work on day 2, and detection day 3. Two microarray chips are used per sample, and the entire process takes 2 and a half days total to complete.

Only normal/non-tumor samples are clustered, and these samples are used to establish the average signal for two copies in a given cluster. The tumor sample signal is then compared to the normal sample signal, and through an iterative process, the copy number for a specific marker is called. To account for saturation at a high copy number, a Langmuir regression is used during analysis. For homozygous clusters, only the signal for the relevant allele is considered and the other allele signal is ignored. For heterozygous clusters, both signals are analyzed and combined together for total copy number analysis and allelic ratio calculation (9,19).

The analysis algorithm chooses the best matching reference or normal samples for a given tumor sample. This is conducted by comparing each tumor sample against each normal reference sample and measuring the Pearson correlation coefficient. The most highly correlated normal samples are used during analysis to generate the copy number data for that tumor. This technique allows for the analysis of tumor samples when a normal pair is not available; however, ideal references are still patient-matched normal samples when available. The hierarchy of relative importance of criteria for selecting normal samples is as follows: 1) Ideal—samples matched closely with tumor samples with respect to (i) preservation, (ii) extraction methods, (iii) specimen age, (iv) tissue, and (v) patient; 2) Next best criteria—samples matched in (i) preservation, (ii) extraction method, (iii) specimen age, and (iv) tissue type; 3) Minimal criteria—samples matched in (i) preservation, (ii) extraction method, and (iii) age.

By calculating the copy number for each allele, the MIP assay technology becomes a powerful tool for examining the allelic ratio and LOH events as discussed previously. The allele ratio is defined as copy B/copy A, where copy B is the minor allele (12). Therefore the allelic ratio falls between 1 and 0. In a normal sample, the allele ratio should be either 0 (homozygous) or 1 (heterozygous). Thus, the allelic ratio data can identify LOH events (including copy-neutral events), and from this, the level of normal tissue or subclonal contamination in a sample can be easily calculated. This is useful for cancer analysis because, as discussed previously, normal contamination often will occur in a typical cancer biopsy and thus may make copy number interpretation difficult. When this is the case, the allelic ratio falls between 0 and 1, and the percent of contamination can be determined (see Figure 2).

Somatic mutation data analysis

In addition to the copy number data, somatic mutation calls are also delivered. These files include the following information: 1) a somatic mutation score for each mutation; 2) a file that provides performance metrics for each assay; and 3) a value that assesses the quality of the sample, defined as median absolute pairwise difference (MAPD) (see Median absolute pairwise difference (MAPD) metric discussed next). The somatic mutation score is a measurement of how far away from the normal wild-type cluster the sample falls. This score is similar to the commonly used Z-score in statistics; however the MAPD value is weighted in the calculation to generate the final somatic mutation score. The larger the score, the farther from the wild-type cluster the sample falls

and the more likely the somatic mutation is real. The somatic mutation data is the newest feature of the MIP assay, and researchers continue to validate its accuracy and validity. The ability to interrogate the commonly acquired cancer mutations would be a unique and important feature of this novel SNP array.

Median absolute pairwise difference (MAPD) metric

The main metric of sample performance for the MIP assay is median absolute pairwise difference (MAPD). This value is calculated by sampling the copy number for 10,000 markers randomly across the genome. By selecting non-adjacent markers across the genome, this value becomes independent of biological information, thereby allowing it to be a robust measurement of the experimental noise for both tumors and normal samples. Figure 2 in reference (12) shows different samples with varying MAPD values. Note that a linear relationship exists between 2p-RSE and MAPD. MAPD $=$ 2p-RSE $\times$ 1.4. A MAPD value of 0.6 or lower generally indicates acceptable sample quality. As the MAPD values increase, so does the overall noise in the copy number data. Although a higher MAPD does indicate more noise in the sample, it does not necessarily imply that copy number or somatic mutation cannot be detected on the sample.

Copy number and somatic mutation data deliverables

Upon completion of the assay and data analysis, data is compiled and delivered back to the researcher. The final data deliverables include: 1) somatic mutation calls for samples that pass the quality criteria; 2) copy A, copy B, total copy number, allelic ratio defined as Copy B/Copy A (where Copy B is the minor allele), and allelic difference defined as (Copy A $-$ Copy B)/(Copy A $+$ Copy B); and 3) a formatted file for Nexus Copy Number software to visualize total copy number data and allelic tracks, enabling LOH detection. With the Nexus Copy Number software program, users can pre-process data to remove systematic artifacts and reduce "noise" (e.g., GC wave correction); correlate copy number with gene expression data (when available); compare different sample populations to identify genomic regions with statistically significant differences; perform multi-sample analysis to correlate copy number and somatic mutation data with clinical outcomes (when available), including survival analysis to identify regions that can be predictive of outcome; and identify common recurrent aberrations using the GISTIC algorithm. Currently, the OncoScan FFPE Express 2.0 Service offering includes a Nexus Copy Number software demo license and a live data review session with BioDiscovery's analysis service scientists. See Figure 6 for an example of MIP data analysis and visualization.

Publications using MIP assay technology on clinical samples

Since its recent introduction, there have been over 30 publications citing the use of MIP assay technology, with most studies related to cancer. We will focus our review on the published articles and recent presentations from national meetings that describe the application of this novel MIP assay platform for cancer copy number alterations (CNAs) across the genome and somatic mutation detection. See Table 1 for a summary of these clinical publications.

Colorectal cancer

The very first publication detailing the use of MIP assay technology on clinical samples was on colorectal cancer (CRC) (20). In this investigation, the authors designed an initial set of probes (which eventually was used as the basis for the current OncoScan™ FFPE Express 2.0 Service) to interrogate several hundred individual exons of $>$200 cancer genes throughout the genome, and also custom-designed $>$100 probes in close proximity of microsatellite markers on chromosome arm 18q, which is known to be prognostically significant in CRC. Of note, the initial number of probes on this first MIP assay was only 972, which was reduced to a 667 MIP assay panel for the final analysis. The authors studied 18 CRC cell lines and 33 primary CRC samples from patients with stage II and III disease. Using the MIP assay data, the investigators distinguished different categories of genomic instability among both the cell lines and primary CRC samples, including specific patterns of 18q deletions. Of note, one of the co-authors of this study was Dr. Ronald Davis, director of the Stanford Genome Technology Center at Stanford University where the MIP assay technology was developed before the founding of ParAllele Bioscience.

In a recent study presented at the American Society of Clinical Onoclogy (ASCO) 2012 Gastrointestinal Cancers Symposium (San Francisco, CA), the MIP assay was again used in CRC where genomic differences between primary tumors and metastatic nodes were examined (21). The 335K MIP cancer panel (Oncoscan™ FFPE Express 2.0 Service) was used to analyze copy number changes in 30 primary and lymph node metastasis CRC pairs (dissected for $>$80% tumor from paraffin-embedded tissue). The data demonstrated significant differences between the primary tumor and lymph node metastasis in four chromosomal regions, including 7p11.2 (containing *EGFR*), 8q24, 14q11.2, and 20q12-q13. In all cases, the metastatic lymph nodes were copy number$-$normal compared to the copy number gains observed in the primary tumor.

In a third CRC study, researchers at the Institut Català d'Oncologia in Barcelona, Spain studied CRC tumors that were due to unknown familial origin (22). Characteristics often associated with Lynch syndrome—abnormal mismatch repair processes and microsatellite instability were absent in these samples and, therefore, the genetic origin of the tumors were unknown. These inherited mutations are grouped together as familial colorectal cancer of type X (fCRC-X). A 330K MIP assay cancer panel (Oncoscan FFPE Express Service) was used to study copy number aberrations and copy number$-$neutral LOH in 16 fCRC-X tissues. This experiment led to the discovery of a unique representative karyotype of fCRC-X tumors that is now being investigated further.

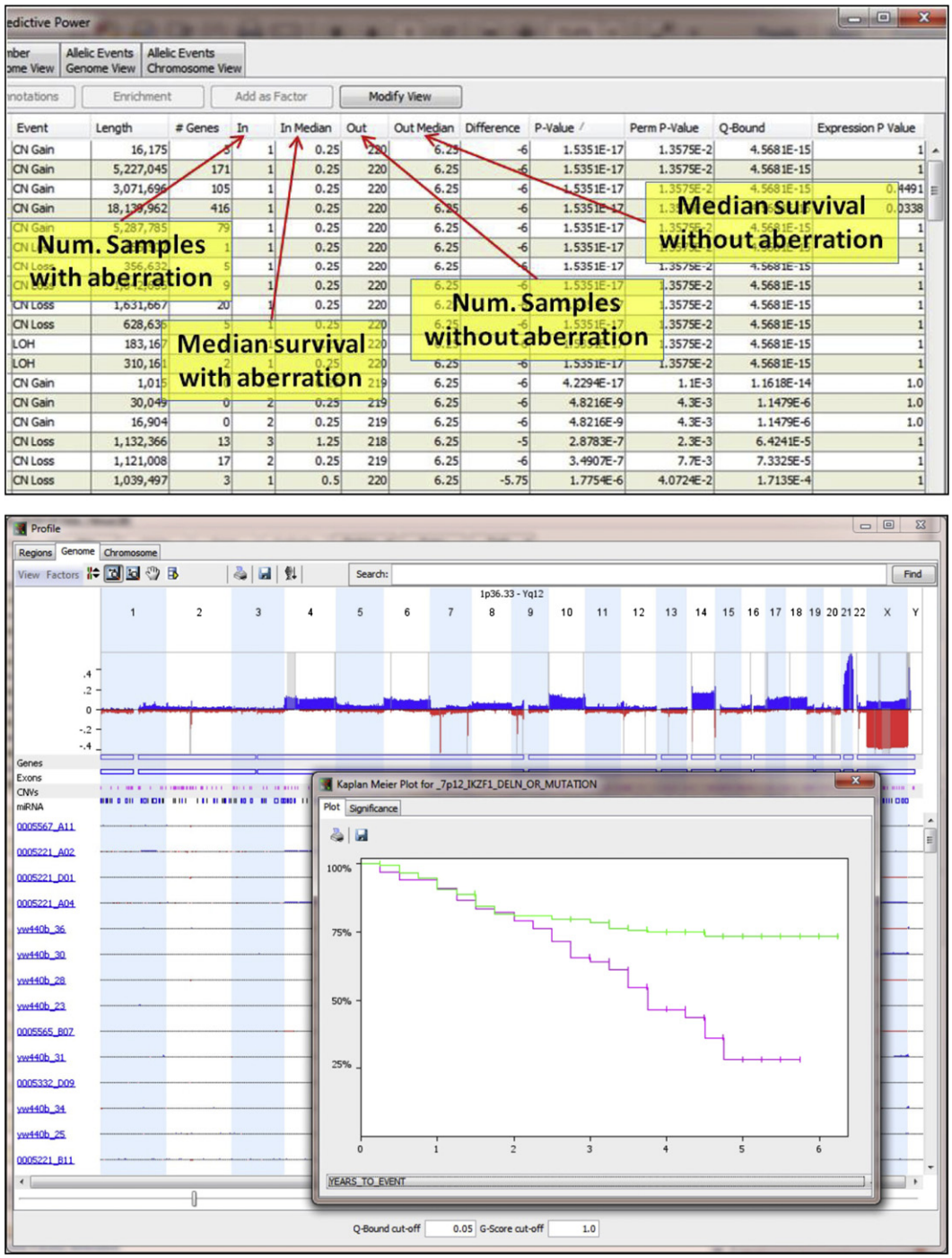

Figure 6 Data analysis with Nexus Copy Number software allows for easy manipulation and analysis of the data, including the creation of relevant data columns with clinical information as highlighted in the top part of the figure. The bottom figure shows a GISTIC analysis identifying significant recurrent aberrations in a population of 170 samples, and the survival plot shows correlation of one event with outcome (BioDiscovery, Inc., El Segundo, CA).

Childhood leukemia

Another early MIP study used clinical samples focused on childhood leukemia (23). Childhood leukemia accounts for >30% of newly diagnosed childhood malignancies, and these investigators performed a pilot study with the MIP platform on freshly collected bone marrow aspirates from 45 primary pediatric leukemia patients (precursor B-cell acute lymphoblastic leukemia ([ALL] [n = 23], precursor T-cell ALL [n = 6], and acute myeloid leukemia [n = 14]). The MIP assay consisted of an extended cancer panel based on the initial CRC study, which included 24,037 SNPs still focused mostly on cancer genes. This MIP assay analysis identified 69 regions of recurring copy number changes, of which 41

Table 1 Summary of MIP publications using clinical samples

Sample/disease type	Year published	MIP panel (no. of probes)	Significant findings	Reference
Colorectal cancer (CRC)	2006	667	Distinguished different categories of genomic instability among CRC cell lines and primary samples, including patterns of 18q deletions.	(20)
Childhood leukemia	2009	24K	Identified 69 regions of recurring copy number changes, including structural pattern of 9p21.3 (*CDKN2A*) copy number loss in precursor B-cell ALL vs. precursor T-cell ALL.	(23)
Ovarian cancer	2008	53K	Classified high-grade serous ovarian carcinomas into three genomic sub-types that display unique molecular signatures involving the PI3K/AKT and p53 pathways.	(28)
Ovarian cancer	2008	51 genes	Confirmed the majority of recurrent 11q13 amplifications in serous ovarian tumors map to four candidate oncogenes and a second more telomeric amplicon with no prior known candidate genes.	(29)
Ovarian cancer	2011	330K	Identified a high percentage of CNAs in both benign and borderline serous tumors, which were found in fibroblasts, suggesting that benign serous tumors may in fact be primary fibromas.	(30)
Ovarian cancer	2011	330K	Correlated *FOXL2* mutation status and genomic imbalances in adult granulosa cell tumors (AGCTs) and provided further details of the mechanisms for mutant allelic imbalance.	(31)
Breast cancer	2011	330K	Identified CNA signatures that predicted recurrence risk among early-stage breast tumors, independent of estrogen receptor status. Also described specific CNAs involved in either increasing or decreasing risk of distant metastasis.	(33)
Breast cancer	2011	330K	Additional 20% survival advantage explained by CNAs associated with tumor detection method. CNA signatures may help in choice of surgical and/or adjuvant treatment to determine indolent vs. aggressive screen-detected tumors.	(34)
Breast cancer	2011	330K	Identified genes that promote the ductal carcinoma in situ (DCIS) phenotype, found evidence that DCIS continues to evolve in parallel with co-existing invasive ductal carcinoma (IDC), and identified potential therapeutic targets.	(35)
Breast cancer	2011	330K	Identified that human ductal carcinoma in situ (DCIS) subtype-specific pathology is preserved in a mouse intraductal (MIND) xenograft model.	(36)
Breast cancer	2012	330K	Introduced "number of subchromosomal regions with allelic imbalance extending to the telomere (NtAI)" to predict initial response to platinum-based chemotherapy in serous ovarian cancer samples without *BRCA1/2* mutations.	(37)
Melanoma	2012	330K	Demonstrated the ability to distinguish benign pagetoid Spitz nevi with no genomic instability/mutations from malignant melanoma lesions.	(38)
Melanoma	2012 (in press)	330K	Described use of MIP assay to detect malignant melanoma with unambiguous histology with 89% sensitivity and 100% specificity (as well as distiniguish those with less-clear histology).	(39)
Bone cyst	2011	330K	Reported unusual *SS18* deletion in solid aneurysmal bone cyst (ABC) and the detection of over 10 regions of CNAs and LOH that were not found by standard cytogenetic analysis.	(40)
Burkitt lymphoma	2011	330K	Performed genome-wide analysis of Burkitt lymphoma in FFPE to find both known and novel recurrent CNAs and LOH, including those that correlate with outcome.	(41)
Pediatric astrocytomas	2010	330K	Characterized new subset of pediatric malignant astrocytomas with combined $BRAF^{V600E}$ mutation and *CDKN2A* deletions, and that total number of CNAs increase with tumor grade.	(42)
Ewing sarcoma	2012	330K	Described novel and known CNAs in Ewing sarcoma, including CNAs unique to relapsed samples. Also proposed multifactor copy number (MCN)-index signature based on eight CNA loci that correlates with clinical outcome.	(43)

had not yet been identified using other DNA microarray platforms. The CNAs were validated in >98% of clinical karyotypes and available fluorescence in situ hybridization (FISH) reports. The authors also described a unique structural pattern of 9p21.3 (*CDKN2A*) copy number loss detected by the MIP assay in precursor B-cell ALL patients versus precursor T-cell ALL patients.

The same authors subsequently transitioned to the 330,000 MIP assay cancer panel (OncoScan FFPE Express Service). They have demonstrated the robustness of this assay on clinically archived FFPE bone marrow aspirate clots from diagnostic precursor B-cell ALL samples stored in FFPE from 1 to 3 years (N = 60) (24). The investigators have since used these same FFPE samples to describe novel deletions in precursor B-cell ALL, which may be related to B-cell development (25), as well as identified specific genomic lesions that may serve as molecular prognostic markers for clinical outcome in childhood ALL (26,27).

Ovarian cancer

The MIP assay has been instrumental in improving the understanding of tumorigenesis processes in ovarian cancers and stages of disease progression. Two of the other original publications to describe the use of the MIP assay with clinical samples were based on ovarian tumors. The first study used MIP assay technology to aid in the genetic profiling of 35 high-grade serous/undifferentiated ovarian tumors (28). For their investigation, the authors of this first MIP assay technology and ovarian cancer study wanted to better understand the molecular mechanisms associated with *BRCA1/2* loss as a way to further sub-classify high-grade serous carcinomas. They began with 49 overall ovarian cancers and first looked at *BRCA1/2* mutations status, *BRCA1* LOH, *BRCA1* promoter methylation, *BRCA1* gene expression, and BRCA1 immunohistochemistry. They subsequently followed this up with MIP assay analysis; *PIK3CA* and *PTEN* expression by qRT-PCR; and p21, p53, and WT-1 immunohistochemistry on 35 of the 49 high-grade serous/undifferentiated tumor samples without *BRCA2* mutations. For the MIP assay analysis, the study investigators used the then-available 53K MIP assay panel and focused on the *MYC* and *PIK3CA* loci. In their analysis, they found decreased *PTEN* mRNA levels in *BRCA1*-mutated, high-grade serous ovarian carcinomas compared to *PIK3CA* copy number gains in those ovarian tumors with *BRCA1* epigenetic loss. Furthermore, those tumors with *BRCA1* epigenetic loss were also more likely to overexpress p53 with concomitant p21 loss. The authors concluded that high-grade serous ovarian carcinomas could be classified into three groups (*BRCA1* loss [genetic], *BRCA1* loss [epigenetic], and no *BRCA1* loss) that display unique molecular signatures involving the PI3K/AKT and p53 pathways. This paper also illustrated early on that MIP assay technology could be combined with other molecular techniques to help categorize tumors into distinct genomic subtypes.

Some of the same authors then continued this work in ovarian carcinoma by focusing their efforts on better understanding the recurrent 11q13 amplification observed in these tumors, as well as in breast, head and neck, oral, and esophageal cancer (29). The investigators studied four candidate oncogenes (*EMSY*, *PAK1*, *RSF1*, and *GAB2*) within this 11q13 locus in 538 clinically annotated ovarian carcinomas with 12 years of follow-up data using standard FISH technology. As part of this study, they used a custom-designed MIP assay panel for 51 genes located within the entire 11q13-14 region in 33 of the high-grade serous ovarian carcinomas. They discovered that most of the amplification mapped to their four candidate oncogenes; however, they also found a second, more telomeric amplicon with no prior known candidate genes. Here, the MIP assay technology helped the authors to conclude that the 11q13 amplicon in ovarian cancer is likely to be important due to several different genes instead of a single oncogene.

In a more recent study, investigators determined the precursors of serous ovarian carcinomas by analyzing underlying genomic changes that may indicate early events in tumorigenesis (30). High-resolution copy number analysis was conducted with both the newer 330K MIP assay panel (OncoScan FFPE Express Service) and the SNP 6.0 (Affymetrix Genome-Wide Human SNP Array 6.0) microarray on DNA from the epithelial and fibroblast components of a cohort of benign (N = 39) and borderline (N = 24) serous ovarian tumors. All the samples included in the microarray analysis were fresh-frozen. The authors found unexpected results: CNAs were found in fibroblasts from 13 of 39 (33%) of the benign serous tumors and in 3 of 20 (15%) of the serous borderline tumors. Furthermore, 12 of these 16 total fibroblast samples with CNAs contained a gain in chromosome 12. This study, which included the newer MIP assay panel, was one of the first high-resolution genomic analyses of benign serous ovarian tumors. Chromosome 12 trisomy previously had been identified in pure fibromas, and so the authors concluded, based on their findings, that a significant proportion of benign serous tumors may in fact be primary fibromas.

A fourth published MIP assay study in ovarian cancer examined *FOXL2* mutations and large-scale genomic imbalances in FFPE clinically archived samples from adult granulosa cell tumors of the ovary (N = 21) (31). Adult granulosa cell tumors (AGCTs) are a rare class of ovarian tumors with recurrent cytogenetic abnormalities, including *FOXL2* C134W mutant allele imbalance, although the mechanism for allelic imbalance was not previously described. In this study, DNA was extracted from 21 AGCT FFPE tissue and two experiments were run in parallel on the *FOXL2* gene: DNA gene sequencing and microarray analysis with the MIP 330K assay panel (OncoScan FFPE Express Service). The MIP assay detected recurrent genomic imbalances of 14q gain, 16q loss, and 22q loss. Pyrosequencing detected *FOXL2* mutant allele imbalance in three tumors, and the MIP assay found a 32.5 Mb deletion in one case and a 70.9 Mb stretch of homozygosity in a second case that both encompassed *FOXL2*. The mutations identified by pyrosequencing were all consistent with the allele ratios detected by the MIP assay. This study provided the first correlation of *FOXL2* mutation status and genomic imbalances in AGCTs, and it further elucidated the mechanisms for mutant allele imbalance in cancer.

A study recently presented at the 2012 AACR Annual Meeting described the use of MIP assay technology to identify biomarkers for the effectiveness of treatment in an ovarian cancer phase 1 clinical trial (32). Specifically, these

investigators from Genentech, Inc. (South San Francisco, CA) studied selective PI3K and dual PI3K/mTOR inhibitors to confirm pathway inhibition at tolerable doses and to identify predictive and pharmacodynamic biomarker candidates. Working on 150 archived ovarian tumors, they employed various laboratory methods, including a six-gene mutation assay (*PIK3CA*, *EGFR*, *KRAS*, *BRAF*, *NRAS*, *AKT1*), PTEN immunohistochemistry, FISH for PIK3CA, and the 335K MIP assay panel (OncoScan FFPE Express 2.0 Service) on select samples for an expanded mutation panel and genome-wide copy number alterations. The authors describe overall 7% *PIK3CA* mutations, 15% PTEN loss, and several samples with *PIK3CA* amplification, and that two of five confirmed responses were in patients with *PIK3CA*-activating mutations. These results led the authors to conclude that patient stratification in phase II clinical studies will be useful to determine if predictive biomarkers can identify those patients who will respond to therapy. The ability of the MIP assay to perform on FFPE archived samples may make it a useful tool for similar use in other previous and ongoing clinical trials.

Breast cancer

Analysis of copy number imbalances has been an effective tool in understanding the underlying molecular mechanisms for breast cancer, and the MIP assay technology has recently contributed to these studies. In fact, Thompson et al. performed one of the largest genomic studies using the MIP assay technology on FFPE stage I/II breast cancer samples from 971 women (33). The aim of this study was to determine copy number gains and losses that could be used clinically for breast cancer. Twelve novel CNAs were found to correspond with clinical prognosis: losses at 1p12, 12q13.13, 13q12.3, 22q11, and Xp21; and gains at 2p11.1, 3q13.12, 10p11.21, 10q23.1, 11p15, 14q13.2-q13.3, and 17q21.33. Using an additional seven previously known CNAs associated with outcome, the authors were able to identify CNA signatures that that significantly predicted recurrence risk among early-stage breast tumors, independent of estrogen receptor status. The MIP assay results also revealed evidence for specific CNAs involved in either increasing or decreasing the risk of distant metastasis in breast cancer.

The same group followed up with a second study on breast cancer to test if screening mammography−detected tumors and symptom-detected tumors would show genotypic differences as copy number alterations (34). Taking a similar approach as described previously, the authors analyzed FFPE breast cancer samples from 850 women with the same MIP 330K cancer panel. Interestingly, they found that chromosomes 2p, 3q, 8q, 11p, and 20q gains were associated with the method of breast cancer detection ($P < 0.00001$). Moreover, in addition to the standard markers of breast tumor aggressiveness, an additional 20% survival advantage could be explained by the CNAs associated with the detection method. These findings may help in the choice of surgical and/or adjuvant treatment for breast cancer by the ability to now use these CNA signatures to determine indolent versus aggressive screen-detected tumors.

Ductal carcinoma in situ (DCIS) is a precursor to invasive ductal carcinoma (IDC), and another breast cancer study used MIP assay technology to identify the genes that promote this phenotype (35). The authors microdissected DCIS samples with synchronous IDC ($N = 21$) and matching normal controls, and genomic DNA was extracted and run on the original 330K MIP assay panel (OncoScan FFPE Express Service Assay). Matched IDC and DCIS samples contained very similar copy number profiles, with an average 83% of shared genome. Findings also included IDC-specific recurrent 3q, 6q, 8p, and 11q loss; 5q, 16p, 19q, and 20 gains; and DCIS-specific 17p11.2 loss. *CCND1* and *MYC* were amplified in the IDC samples, along with *AXL*, *SPHK1*, and *PLAUR* (all previously associated with aggressive breast cancer and possible candidate genes for therapeutic intervention). Using MIPs on these FFPE archived samples, the authors provided evidence that the DCIS continues to evolve in parallel with the co-existing IDC as well as identified potential therapeutic targets.

Another group studying breast cancer has created mouse DCIS xenograft models that recapitulate the pathology and heterogeneity of human disease and are now going forward with MIP assay technology, as outlined in their recent paper, to confirm that the cells from the mice are truly DCIS on a molecular level and not hyperplastic or normal human cells (36).

One of the most recent breast cancer studies to use MIP assay technology studied genomic signatures indicative of defective DNA repair in cell lines and tumor samples (37). In addition to 10 breast cancer cell lines, the investigators obtained 27 and 28 triple-negative breast cancer core biopsies from two clinical trials using cisplatin (Cisplatin-1) and cisplatin+bevacizumab (Cisplatin-2), respectively. Most of the samples were fresh-frozen specimens, although several FFPE samples were also included. Labeling, hybridization, and data processing was performed first with the 70K MIP assay panel and then subsequently the 330K MIP assay panel when they became available. This study found that cisplatin sensitivity in vitro and preoperative pathologic response was predicted by the number of subchromosomal regions with allelic imbalance extending to the telomere (N_{tAI}), irrespective of *BRCA1* and *BRCA2* status. The authors then validated the ability of N_{tAI} to predict initial response to platinum-based chemotherapy in serous ovarian cancer samples without *BRCA1/2* mutations from The Cancer Genome Atlas Project (TCGA). This study concluded that patients without *BRCA1/2* mutations may also benefit from DNA damaging agents such as cisplatinum, especially when the N_{tAI} is considered. Using the MIP assay, it may become feasible to clinically measure N_{tAI} status in preoperative breast cancer samples in FFPE with low tumor cell counts.

Melanoma

Two publications have described the use of MIP assay technology related to cutaneous oncology, particularly melanoma. One of the ongoing challenges in the clinical field of dermatopathology is the ability to distinguish benign from malignant lesions. In the first report, genomic analysis was used to understand the genetic mechanisms and the malignant potential of pagetoid Spitz nevi by investigating correlations between copy number alterations and the disease (38). Pagetoid spitz lesions cannot be easily differentiated from melanoma, and so the researchers were interested in

using molecular technology to better study this type of lesion. Several experiments were performed on limited FFPE pagetoid Spitz nevi and normal tissue biopsies from a single patient at diagnosis, including germline karyotyping; customized aCGH (ExonArrayDx, version 1.0; GeneDx) for melanoma-associated genes (*HRAS*, *BRAF*, *KRAS*, *RAF1*, *CDKN2A*, *Rb1*, *MAP2K1*, *MAP2K2*, *PTEN*, and *PTPN11*); gene sequencing for *BRAF*, *CDK4*, and *CDKN2A*; and genome-wide copy number and LOH analysis using the 330K MIP assay panel (OncoScan™ FFPE Express Service). No genomic instability and no mutations were found, which helped to confirm this difficult-to-diagnose lesion as non-melanoma and the suggestion that such types of genomic analyses may be useful for the identification of melanoma mimics.

Likewise, a second investigation studied the use of the MIP assay to distinguish malignant melanomas from benign nevi (39). Despite the effective the use of histology in the clinical setting to detect malignant melanoma, a limited number of biopsies reviewed by dermatopathologists will remain uncertain as to their malignant potential; however, this type of information is required for choosing the optimal treatment course for patients with suspicious skin lesions. The authors used the 330K MIP assay panel (OncoScan FFPE Express Service) on FFPE samples to differentiate benign nevi ($n = 23$) from malignant melanoma ($n = 30$) and to predict the clinical course of a set of histologically ambiguous melanocytic lesions ($n = 11$). The authors found that the MIP assay could detect malignant melanoma with unambiguous histology with 89% sensitivity and 100% specificity, but did not perform as well for cases with less-clear histology. Nevertheless, the performance of the assay was impressive. This study demonstrated that genome-wide analysis for CNAs and LOH in melanocytic lesions with arrays such as MIP technology increased the likelihood of detecting pathogenic alterations, and more so than the currently available targeted FISH assays. The authors appropriately caution, however, that widespread clinical adoption of novel microarray technology will require appropriate clinical follow-up and large validation sets.

Prostate cancer

MIP assay technology also can be used as a reliable platform to validate whole genome sequencing results from cancer studies. In a recent study presented at the 2012 AACR Annual Meeting, investigators used MIP assay technology (OncoScan FFPE Express 2.0 Service) to validate whole genome sequencing results from 50 prostate cancer biopsies with limited DNA by the Canadian Prostate Cancer Genome Network (CPC-GENE). Sequencing was initially performed on the Illumina HiSeq 2000 platform (Illumina, San Diego, CA) to generate coverage depths of 50× for tumor samples and 30× for reference samples. Following sequence alignment and variant-calling, the results were compared to the assay results. Single-nucleotide variants detected using arrays were validated >99% of the time by the sequence data. Furthermore, the copy number variant (CNV) calls made by the MIP assay also matched the CNV calls. The authors were able to conclude that their sequencing did not exhibit genome-wide coverage biases, and now will proceed to sequence an additional 450 prostate samples that may also contain limited amounts of initial start-up DNA. As more sequencing is attempted in other cancers, the MIP assay technology may continue to serve a useful role in genotype and structural variation validation given its ability to also perform well with low amounts of input DNA.

Bone cyst

Another case report using MIP assay technology helped to demonstrate an unusual deletion in *SS18* in a solid aneurysmal bone cyst (ABC) (40). Oftentimes, ABCs will contain *USP6* rearrangements, but *SS18* rearrangement will be limited to the diagnosis of synovial sarcoma. In the patient described in this article, standard FISH analysis indicated the unusual finding of rearrangements of both *USP6* and *SS18*, but all the histologic features still remained consistent with a diagnosis of ABC. The clinical karyotype revealed 45,X,add(X)(p11.2),add(4)(q13),add(8)(p21),-13,add(17)(p11.2),add(18)(q11.2) in all 20 cells analyzed, with no evidence for *USP6* rearrangement. Through a series of follow-up analyses by RT-PCR, immunohistochemistry, and the 330K MIP assay panel (OncoScan™ FFPE Express Service), the investigators were able to rule out *SS18-SSX1*, *SS18-SSX2*, and *SS18-USP6* translocations. Refining this analysis further, the MIP assay demonstrated a deletion encompassing the entire *SS18* gene and its promoter, as well as regions targeted by the commercial *SS18* FISH probe (which may help to explain the unusual FISH results). In addition, over 10 regions of CNAs and LOH were detected by the MIP assay that were not found in the standard cytogenetic analysis (although some of these LOH regions probably represent constitutional changes). The authors concluded that commercially available FISH probes may occasionally yield misleading results, which can be clarified by molecular analysis. This case report is a good example of the ability of MIP assay technology to assist in the work-up of unusual cases where no fresh-frozen tissue is banked prospectively, but FFPE tissue is readily available for analysis.

Burkitt lymphoma

Genetic markers for high-risk pediatric Burkitt lymphoma (pBL) are unknown. In one of the first studies of genome-wide copy number changes by high-resolution microarray in pBL, the authors used the MIP assay on clinically archived lymphoma tissue specimens to identify both known and novel recurrent copy number abnormalities (41). Although the majority of patients with pBL will be cured, many of those who relapse will die of disease, and so the authors sought to identify genetic markers that correlate with clinical outcome. Similar to the other studies previously described, the investigator used the 330K MIP assay panel (OncoScan FFPE Express Service) to analyze 28 clinically archived FFPE tumor specimens and 24 paired FFPE normal specimens. Important for their study, nearly every sample could be linked to follow-up data so that clinical outcome could be assessed and correlated with genomic loci. The authors identified CNAs in 64% of samples, with 32 specific gains and 30 losses. Seven recurrent CNAs were found and included 1q gain (25%), 13q gain (11%), and 17p loss (14%). The study

reported on a minimum common amplified region at 13q31 that included the *MIR17HG* (*MIR17-92*) locus. Furthermore, samples with this focal gain had higher *MIR17* RNA expression and were more likely to relapse. The authors also reported that uniparental disomy (copy-neutral LOH) could be identified in 32% of cases and was often recurrent. This study represents an excellent example of how FFPE samples with outcome data can be useful when analyzed by MIP assay to identify loci that may have clinical prognostic value as well as discover novel therapeutic targets.

Pediatric astrocytoma

The underlying genetic basis of malignant astrocytomas is not clearly understood. Investigators used the earlier version of the MIP assay cancer panel consisting of 24,037 SNPs to identify CNAs associated with different grades of astrocytomas in children (42). The investigators interrogated 33 primary diagnostic specimens that were histological grade 1−4 tumors. Interestingly, gains of 10-fold or more were found only in the higher grade 3−4 tumors, which have a worse clinical outcome, and included *MDM4* (1q32), *PDGFRA* (4q12), *MET* (7q21), *MYC* (8q24), *PVT1* (8q24), *WNT5B* (12p13), and *IGF1R* (15q26). Homozygous deletions of *CDKN2A* (9p21), *PTEN* (10q26), and *TP53* (17p3.1) were found in the grade 2−4 tumors. The fraction of the genome altered by CNAs increased in a statistically significant manner as the tumor grade progressed. Using the MIP assay, the study found a recurrent 2 Mb amplification of 7q34 including *BRAF* in the low-grade tumors. This prompted further evaluation of this locus by RT-PCR and DNA sequencing, and the *KIAA1549-BRAF* fusion transcript was found to be expressed in 10 of 10 (100%) of the grade 1 astrocytomas (and none of the grade 2−4 tumors) and the oncogenic missense $BRAF^{V600E}$ mutation was detected in 7 of 31 (23%) of the grade 2−4 tumors (but none of the grade 1 tumors). Upon closer inspection, the authors discovered that *CDKN2A* deletions occur quite frequently in cases with $BRAF^{V600E}$ mutations (N = 5 of 7, 71%), perhaps defining a new subset of pediatric astrocytomas.

Ewing sarcoma

One of the most recent studies to use the MIP assay interrogated clinically archived FFPE samples from patients with Ewing sarcoma (43). Ewing sarcoma is the second most common bone tumor in children and young adults with a very poor clinical outcome for metastatic and relapsed disease, but limited studies have described molecular CNAs in this disease. Jahromi et al. (43) used the 330K MIP assay panel (OncoScan FFPE Express Service) to investigate FFPE specimens from Ewing sarcoma patients in 40 primary tumors and 12 metastatic lesions. They identified homozygous deletions in *SMARCB1* (*INI1/SNF5*) in an unusual subset of atypical-appearing Ewing sarcoma tumors. They also correlated specific CNAs with clinical outcome, proposing a multifactor copy number (MCN)-index based on eight different CNA regions that was highly predictive of overall survival (39% vs. 100%, $P < 0.001$). Finally, analyzing the metastatic samples, they identified *RELN* gene deletions that were unique to 25% of Ewing sarcoma metastatic samples. Using the MIP assay, they were able to describe both known and novel CNAs in FFPE tumor blocks, up to 12 years of age, from patients previously diagnosed with Ewing sarcoma. The authors proposed their CNA prognostic signature based on the mature clinical data associated with the clinically archived FFPE specimens, which is another advantage of using a technology that works on FFPE.

In summary, the development of the MIP assay technology has allowed the field of cancer clinical research to make important new discoveries by enabling accurate copy number, LOH, and somatic mutation analysis in a variety of different cancer types. One of the advantages of MIP technology includes the ability to obtain high-quality results from very small starting quantities of genomic DNA in the amount of 75 ng (and sometimes less). Moreover, due to the design of the MIP assay, it performs remarkably well in highly degraded tissue samples from FFPE blocks. This is very important in cancer research because nearly all of these FFPE samples will have diagnostic and clinical information that can be correlated to the genomic findings detected by the MIP assay. As researchers continue to explore the genetic basis of cancer, it is likely that MIP technology will remain important in the identification of diagnostic biomarkers, the detection of prognostic signatures, and the discovery of candidate genes for developmental therapeutics.

References

1. Nilsson M, Malmgren H, Samiotaki M, et al. Padlock probes: circularizing oligonucleotides for localized DNA detection. Science 1994;265:2085−2088.
2. Hardenbol P, Baner J, Jain M, et al. Multiplexed genotyping with sequence-tagged molecular inversion probes. Nat Biotechnol 2003;21:673−678.
3. Frazer KA, Ballinger DG, Cox DR, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007; 449:851−861.
4. Daly TM, Dumaual CM, Miao X, et al. Multiplex assay for comprehensive genotyping of genes involved in drug metabolism, excretion, and transport. Clin Chem 2007;53:1222−1230.
5. Dumaual C, Miao X, Daly TM, et al. Comprehensive assessment of metabolic enzyme and transporter genes using the Affymetrix Targeted Genotyping System. Pharmacogenomics 2007;8: 293−305.
6. Hu Y, Ehli EA, Nelson K, et al. Genotyping performance between saliva and blood-derived genomic DNAs on the DMET array: a comparison. PLoS One 2012;7:e33968.
7. Biegel JA. Highlights from the second annual Cancer Cytogenomics Microarray Consortium (CCMC) meeting. Cancer Genetics 2011;204:413−415.
8. Pinard R, de Winter A, Sarkis GJ, et al. Assessment of whole genome amplification-induced bias through high-throughput, massively parallel whole genome sequencing. BMC Genomics 2006;7:216.
9. Wang Y, Moorhead M, Karlin-Neumann G, et al. Allele quantification using molecular inversion probes (MIP). Nucleic Acids Res 2005;33:e183.
10. Pinto D, Darvishi K, Shi X, et al. Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants. Nat Biotechnol 2011;29:512−520.
11. Wang Y, Moorhead M, Karlin-Neumann G, et al. Analysis of molecular inversion probe performance for allele copy number determination. Genome Biol 2007;8:R246.
12. Wang Y, Carlton VE, Karlin-Neumann G, et al. High quality copy number and genotype data from FFPE samples using

Molecular Inversion Probe (MIP) microarrays. BMC Med Genomics 2009;2:8.
13. Blow N. Tissue preparation: tissue issues. Nature 2007;448: 959–963.
14. Lapunzina P, Monk D. The consequences of uniparental disomy and copy number neutral loss-of-heterozygosity during human development and cancer. Biol Cell 2011;103:303–317.
15. Maciejewski JP, Tiu RV, O'Keefe C. Application of array-based whole genome scanning technologies as a cytogenetic tool in haematological malignancies. Br J Haematol 2009;146: 479–488.
16. Tuna M, Knuutila S, Mills GB. Uniparental disomy in cancer. Trends Mol Med 2009;15:120–128.
17. Absalan F, Ronaghi M. Molecular inversion probe assay. Methods Mol Biol 2007;396:315–330.
18. Ji H, Welch K. Molecular inversion probe assay for allelic quantitation. Methods Mol Biol 2009;556:67–87.
19. Moorhead M, Hardenbol P, Siddiqui F, et al. Optimal genotype determination in highly multiplexed SNP data. Eur J Hum Genet 2006;14:207–215.
20. Ji H, Kumm J, Zhang M, et al. Molecular inversion probe analysis of gene copy alterations reveals distinct categories of colorectal carcinoma. Cancer Res 2006;66:7910–7919.
21. Tsikitis VL, Larson D, Preece C, et al. Copy number differences in EGFR and other amplicons between primary colon tumors and paired lymph node metastasis. ASCO Meeting Abstracts 2012;30:498.
22. Bellido F, Pineda M, Seguí N, et al. Molecular tumor characterization in familial colorectal cancer of type X. Cellular Oncology 2012;35:37–38.
23. Schiffman JD, Wang Y, McPherson LA, et al. Molecular inversion probes reveal patterns of 9p21 deletion and copy number aberrations in childhood leukemia. Cancer Genet Cytogenet 2009;193:9–18.
24. Schiffman JD, Downie J, Demarest B, et al. Novel use of molecular inversion probes to interrogate formalin-fixed paraffin-embedded (FFPE) samples of childhood leukemia [abstract]. ASH Annual Meeting Abstracts 2009;114:1589.
25. Miles RR, Downie J, Jahromi MS, et al. VPREB1 deletions occur independent of lambda-light chain rearrangement and predict worse outcome in pediatric acute lymphoblastic leukemia (ALL) [abstract]. ASH Annual Meeting Abstracts 2010;116:273.
26. Downie J, Barnette P, Rodic V, et al. T-cell receptor gene deletions are associated with high risk features and worse outcome in childhood precursor B-cell acute lymphoblastic leukemia (ALL) [abstract]. ASH Annual Meeting Abstracts 2010;116:275.
27. Mangum DS, Shams S, Downie JM, et al. Focal 22q11.22 loss combined with IKZF1 alterations predict very poor outcome in childhood acute lymphoblastic leukemia [abstract]. ASH Annual Meeting Abstracts 2011;118:741.
28. Press JZ, De Luca A, Boyd N, et al. Ovarian carcinomas with genetic and epigenetic BRCA1 loss have distinct molecular abnormalities. BMC Cancer 2008;8:17.
29. Brown LA, Kalloger SE, Miller MA, et al. Amplification of 11q13 in ovarian carcinoma. Genes Chromosomes Cancer 2008;47: 481–489.
30. Hunter SM, Anglesio MS, Sharma R, et al. Copy number aberrations in benign serous ovarian tumors: a case for reclassification? Clin Cancer Res 2011;17:7273–7282.
31. Geiersbach KB, Jarboe EA, Jahromi MS, et al. FOXL2 mutation and large-scale genomic imbalances in adult granulosa cell tumors of the ovary. Cancer Genet 2011;204:596–602.
32. Spoerke J, O'Brien C, Wu J, et al. Biomarker evaluation in phase I clinical trials of selective PI3K and PI3K/mTOR inhibitors American Association of Cancer Research (AACR) Annual Meeting. Chicago, IL, 2012.
33. Thompson PA, Brewster AM, Kim-Anh D, et al. Selective genomic copy number imbalances and probability of recurrence in early-stage breast cancer. PLoS One 2011;6:e23543.
34. Brewster AM, Thompson P, Sahin AA, et al. Copy number imbalances between screen- and symptom-detected breast cancers and impact on disease-free survival. Cancer Prev Res (Phila) 2011;4:1609–1616.
35. Johnson CE, Gorringe KL, Thompson ER, et al. Identification of copy number alterations associated with the progression of DCIS to invasive ductal carcinoma. Breast Cancer Res Treat 2011.
36. Valdez KE, Fan F, Smith W, et al. Human primary ductal carcinoma in situ (DCIS) subtype-specific pathology is preserved in a mouse intraductal (MIND) xenograft model. J Pathol 2011;225: 565–573.
37. Birkbak NJ, Wang ZC, Kim J-Y, et al. Telomeric allelic imbalance indicates defective DNA repair and sensitivity to DNA-damaging agents. Cancer Discovery 2012;2:366–375.
38. Harris K, Florell SR, Papenfuss J, et al. Melanoma mimic: a case of multiple pagetoid Spitz nevi. Arch Dermatol 2012;148:370–374.
39. Chandler WM, Rowe LR, Florell SR, et al. Differentiation of malignant melanoma from benign nevus using a novel genomic microarray with low specimen requirements. Arch Pathol Lab Med (In press);136.
40. Geiersbach K, Rector LS, Sederberg M, et al. Unknown partner for USP6 and unusual SS18 rearrangement detected by fluorescence in situ hybridization in a solid aneurysmal bone cyst. Cancer Genet 2011;204:195–202.
41. Schiffman JD, Lorimer PD, Rodic V, et al. Genome wide copy number analysis of paediatric Burkitt lymphoma using formalin-fixed tissues reveals a subset with gain of chromosome 13q and corresponding miRNA over expression. Br J Haematol 2011;155:477–486.
42. Schiffman JD, Hodgson JG, VandenBerg SR, et al. Oncogenic BRAF mutation with CDKN2A inactivation is characteristic of a subset of pediatric malignant astrocytomas. Cancer Res 2010; 70:512–519.
43. Jahromi MS, Putnam AR, Druzgal C, et al. Molecular inversion probe analysis detects novel copy number alterations in Ewing sarcoma. Cancer Genetics 2012;205:391–404.