BeadArray™-based solutions for enabling the promise of pharmacogenomics

Jian-Bing Fan¹, Sean X. Hu^{1,2}, William C. Craumer¹, and David L. Barker¹

¹Illumina, San Diego, CA and ²IMS Management Consulting, Plymouth Meeting, PA, USA

BioTechniques 39:583-588 (October 2005)

doi 10.2144/000112047

A “one-size-fits-all” approach continues to characterize today’s healthcare paradigm. But emergent rules, information, genomics tools, and economics are driving a fundamental and inevitable shift to a more personalized world of medicine. In this new world, the interests of insurers, regulators, suppliers, healthcare providers, and most important, patients, will have converged. The new goal will be the right treatment for the right individual at the right time. In this world, personalized medicine, through pharmacogenomics (PGx), will be the new healthcare paradigm. We will briefly examine healthcare trends and current opportunities for PGx development. We will then demonstrate how microarray technologies—among them bead-based approaches—have emerged as a key enabler for bringing home the promise of PGx.

INDIVIDUAL DIFFERENCES, INDIVIDUAL OUTCOMES

Therapeutic treatments rarely, if ever, work for all patients. Even blockbuster drugs are efficacious in just 40%–60% of the patients to whom they are administered. And many patients suffer adverse drug reactions (ADRs), so much so that ADRs are a leading cause of death among Americans. After all, people are different. Pharmacogenomics (PGx) aims to understand and then take advantage of those differences.

Genetic variation, however, affects far more than drug response. Increasingly, scientists and clinical researchers are studying individual differences at the molecular level to understand how genetic variation directly impacts differential disease predisposition, symptomatic development, and clinical outcomes in addition to drug response.

GENETIC INFORMATION: OPPORTUNITIES AND BENEFITS

DNA and RNA analysis, in particular, are generating molecular-level information that is driving PGx and personalized medicine. PGx-enabling development is occurring in four principal areas.

Molecular Diagnostics

Progress has already been made in this area with the generation of gene expression profiles that provide increasingly specific disease characterization based on molecular phenotypes. Cancer, in particular, has witnessed extensive gene expression research with the intent of providing accurate diagnosis and ensuring the proper therapeutic approach, irrespective of optimal dosage levels.

Earlier disease diagnosis will, of course, enable earlier intervention and better odds of improved outcomes. There’s a bit of a “Catch 22” embedded here, in that individuals will have little reason to get tested, and providers will have little incentive to approve an early test unless there is family history or unless genotyping and related preemptive approaches have indicated a specific disease predisposition. A further need is for Congress to pass pending legislation, already approved in the Senate, which would protect patients from health insurance discrimination on the basis of inherited characteristics.

Stratification Strategies

Although ADR is a leading cause of hospitalization and death, stratification of populations into responders, nonresponders, and negative responders is still conducted on a trial-and-error basis. Retrospective studies, however, are providing important information about the pattern of genetic variation in each population stratum.

Further studies will elucidate key biomarkers associated with drug response and enable stratification before trials begin to optimize outcomes. Stratification strategies have also resulted in the selective reintroduction of orphan drugs and/or the administration of those drugs to new classes of patients with different diseases but related genotypes. With the profusion of new information about single nucleotide polymorphisms (SNPs) and haplotypes, large-scale SNP genotyping studies are gaining momentum to generate actionable information about disease associations and about predicted therapeutic response.

Improved Therapeutics

Array-based technologies are delivering significant information value to therapeutic developers, with benefits that include: (i) generation of higher-quality targets; (ii) provision of data for earlier target screening and validation; (iii) improved cycle times for lead validation and preclinical studies; (iv) toxicological screening for more efficient elimination of poor candidates; and (v) more efficient clinical development programs.

Predisposition/Outcome Management

Still unrealized, the activity with the greatest potential impact may be the ability to predict, as a function of genotype, an individual’s predisposition to disease. New tool sets and information deriving from the International HapMap Project and related research will trigger a literal avalanche of array-based studies designed to understand genetic variation and correlate genetic variation with disease—especially complex diseases with multifactorial and yet-to-be understood genetic etiologies. Also central to personalized medicine will be the ability to project the expected timing of symptomatic development, and then to actively manage the risk and treatment following onset to ensure optimal clinical outcomes.

CHALLENGES TO CHANGE

Notwithstanding the Food and Drug Administration’s (FDA’s) support of PGx and its powerful promise (1,2), Larry Lesko, Director of the Office of Clinical Pharmacology and Biopharmaceuticals at the agency’s Center for Drug Evaluation and Research, cautions against “genotyping instead of genotyping.”

The challenges are clear. Among them is the massive amount of discovery research required to generate and catalog disease expression profiles and to associate genetic variation with disease. More and more PGx/clinical phenotypes are found to have a multifactorial molecular genesis. As a result, larger patient sample pools will have to be studied to discover and validate molecular biomarkers for clinical applications.

An equally significant but hard-to-qualify challenge may be one of inertia. Healthcare participants are often reluctant to change the status quo, since one assumes there will be winners as well as losers in the new world. The benefits must be understood and communicated to help effect change.

The technical hurdles are within reach. To gain higher-confidence information, one could study more genes or more SNP loci across more samples. One could also use multiple genomic applications to attack the same scientific puzzle, integrating for example, genotyping, gene expression, and epigenetic analyses.

Recent developments in microarray technology, along with powerful new assay technologies, have improved throughput, performance, and cost. And access to new information, such as that contributed by the International HapMap Project (3), has enabled array-based solutions with improved interrogative power. As a result, investigators are at last able to design studies and conduct research at an experimental scale sufficient for generating real, credible, and actionable data to drive realization of PGx.

The discussion above illustrates the need for a variety of flexible assay technologies that enable high-throughput genomic research, biomarker discovery, and inexpensive diagnostic tests that will direct patients to therapy that is effective for their individual genetic make up and disease condition, without serious adverse reactions. Illumina (San Diego, CA, USA) has developed microarray platforms and multiplexed assays for genotyping, gene expression, and DNA methylation that provide these needed tools.

ILLUMINA'S CORE BeadArray™ TECHNOLOGY

Illumina produces microarrays in two different formats: (i) the Sentrix® Array Matrix and (ii) the Sentrix BeadChip (Figure 1). The Array Matrix arranges fiber optic bundles, each containing 50,000 5-μm fibers, into an Array of Arrays™ format that is compatible with and can access the wells of a 96-well microtiter plate. The fiber optic bundles in the Array Matrix assembly are polished flat on both ends. On one end, the core of each fiber is etched to form a nanowell that will accept 3-μm silica beads, each of which has been coated with several hundred thousand oligonucleotides of a particular sequence. Up to 1536 bead types self-assemble onto each fiber bundle, giving an average of over 30 beads of each type (4). In the BeadChip format, one to several microarrays are arranged on silicon slides that have been processed by micro-electromechanical systems (MEMS) technology to also have nanowells that support

Figure 1. BeadArray technology. Illumina's Sentrix arrays use either fiber optic bundles or silicon slides as the BeadArray substrate. Several hundred thousand full-length copies of a unique oligonucleotide are covalently attached to each 3-μm silica bead. From 384 to 250,000 bead types are pooled and self-assembled onto an array. On average, 30 copies of each bead type are present on an array.

self assembly of beads. This BeadChip format provides the flexibility to manufacture arrays with from 1536 to 250,000 bead types or even more, as required by the genomic application. In both formats, an average of over 30 beads of each type is maintained—a strategy that provides the accuracy of multiple measurements and statistical power even when using a small number of arrays. Because self-assembly of beads is used in the manufacturing process, each array is decoded to determine the map of bead type positions. A combinatorial series of brief hybridizations and rinses is used that result in a level of accuracy well beyond the requirements of any application (5). This decoding process also provides a quality control measure of the function of each bead that is incorporated into the final bead map. BeadArray technology is the only microarray technology with functional quality control (QC) of every element in the array.

Each assay that is designed for use on Sentrix Arrays results in a fluorescent signal attached to the individual bead types. To read out these signals, Illumina developed the BeadArray Reader. This reader is a confocal, laser scanning fluorescence imager that quantifies the signal on each bead in two-color channels with a dynamic range of over three orders of magnitude (6).

As high-throughput genomic research is used to sift through tens of thousands of genes and hundreds of thousands of variations in the human genome, smaller numbers of biomarkers suitable for diagnostic assays indicating efficacy and safety will be identified. To address the need for affordable PGx tests, Illumina began developing a new, digital BeadArray Technology platform incorporating beads that use holographic identifiers, so that decoding can occur at the same time as fluorescent readout without interference with the analytical signal. This technology will be deployed for research and diagnostic assay panels that require multiplex levels of 10 to a few hundred measurements.

GENOME-WIDE GENOTYPING AND ASSOCIATION STUDIES

The analysis of SNPs in the human genome may offer the key to understanding genetic differences between individuals and disease states, and eventually improve medical treatments by allowing the prediction of genetically related disease risk and

Figure 2. The Infinium assay. The Infinium assay uses the complete complexity of the genome and scales to the number of bead types that can be placed on an array. The assay consists of four parts: 1, whole genome amplification; 2, fragmentation and hybridization; 3, enzymatic single nucleotide polymorphism (SNP) discrimination by allele-specific extension; and 4, fluorescence staining and scanning.

drug response. Linkage disequilibrium-based pharmacogenomic and disease association research requires an accurate and cost-effective means of genotyping hundreds of thousands of SNPs in large sample collections. Despite advances in genotyping, analysis on this scale remains a challenge, with some of the major obstacles being genome complexity and the limited scalability of multiplexing in target-specific PCR-based genotyping assays.

Illumina has developed a microarray-based genotyping approach that overcomes these limitations, effectively enabling unlimited multiplexing of SNP genotyping reactions. A whole-genome amplification and fragmentation strategy is combined with hybridization to 50-mer probe arrays, followed by an allele-specific primer extension assay (7) (Figure 2). The Infinium™ whole-genome genotyping (WGG) assay employs a one-color allele-specific primer extension (ASPE) assay and two bead types per SNP to genotype over 100,000 SNPs in the exon-centric Sentrix Human-1 Genotyping BeadChip product. The approach is unique among genotyping methods, in that essentially the entire genome is present in the sample, and there is no complexity reduction step. Therefore, the number of SNPs that can be read out is limited only by the number of probes on the array and not by multiplexing in the sample preparation. Genotyping quality was evaluated by designing assays to thousands of loci, including a set of SNPs used to validate the technologies involved in the International HapMap Project. The performance of the assay (call rate of 99.5% and concordance of 99.9%) for this set was on par with Illumina's highly accurate GoldenGate® assay, which was used to genotype over 65% of the SNP loci in the HapMap's first phase.

Single base extension (SBE) offers alternative genotyping biochemistry with proven robustness and the potential to effectively double the information content of the BeadChip. Expanding on our Infinium WGG assay, a two-color SBE version was developed that enables analysis of over 250,000 SNPs on a single array by virtue of using only a single bead type per SNP (F. Steemers et al., manuscript submitted). We analyzed the HapMap QC SNPs and found an assay conversion rate of 89.9%, a call rate of 99.86%, reproducibility of 100%, Mendelian Inconsistencies of 0.0023%, and concordance of 99.98%. These results indicate that the SBE assay is a highly robust and accurate genotyping assay that has the same inherent scalability features as the current Infinium WGG assay. Illumina is now developing a 250 K Human-1 Genotyping BeadChip based on HapMap tag SNPs.

The Infinium whole-genome genotyping assay has the following strengths: (i) high quality and robust genotyping (accuracy >99.9%); (ii) effectively unlimited multiplexing; (iii) low amounts of genomic DNA required (750 ng are needed to genotype over 100,000 SNPs); (iv) relatively unconstrained genome-wide SNP selection; (v) low risk of "carryover" contamination, since PCR is not employed; and (vi) walk away automatable assay protocols from sample preparation to readout.

As whole-genome association studies identify regions of interest for particular diseases or drug responses, researchers will need a way to focus in on these regions and analyze large numbers of subjects. We developed a flexible, accurate, and high-throughput SNP genotyping system for genetic analysis in large populations. The GoldenGate assay uses solid-phase allele-specific extension and ligation and multiplexed amplification to genotype up to 1536 SNPs in each sample using the Sentrix Array Matrix format for 96 samples or a BeadChip for 16 samples (4,8,9) (Figure 3). Investigators can generate approximately 300,000 genotypes per day using a BeadStation and minimal other equipment requirements or greater than 1.6 million genotypes per day in a BeadLab environment with robotics-assisted processes. The scalability of the platform allows the user to perform either small pilot studies or large-scale SNP genotyping association studies. Therefore, it brings a new level of flexibility, throughput, and affordability to genetic research. The system has been used for many genetic

studies, including the International HapMap Project (3). In addition, the GoldenGate assay can tolerate a certain degree of DNA degradation. Highly accurate genotyping results have been obtained with whole-genome amplified DNA samples (10), and DNA samples isolated from formalin-fixed paraffin-embedded (FFPE) normal and cancer tissues (Fan et al. unpublished data).

HIGH-THROUGHPUT GENE EXPRESSION PROFILING

Comparison of gene expression patterns in different cell types, developmental stages and disease states should enable the discovery of characteristic gene expression patterns that can be associated with functionally important states. Microarray-based tumor classification (11–13) as well as treatment response and clinical outcome prediction (14–17) have been demonstrated in many cancer types. These studies proved the feasibility of combining large-scale gene expression profiling with classic morphologic and clinical methods of staging and grading cancer for better diagnosis and outcome prediction. Now, the challenges are how to validate the molecular profiles generated from the genome-wide microarray experiments in large clinical sample sets and translate the findings to the clinic (18–25).

Microarray technologies typically require substantial quantities of snap-frozen tissue. Although many institutions are now maintaining frozen tissue banks, which should facilitate gene expression analysis in the future, few of those now have sufficient clinical follow-up data for retrospective studies aimed at correlating gene expression profiles with clinical outcome. The most widely available sample

Figure 3. GoldenGate genotyping scheme (taken from Reference 9). Two allele-specific oligonucleotides (ASOs) and one locus-specific oligonucleotide (LSO) are designed for each single nucleotide polymorphism (SNP). All three oligonucleotide sequences contain regions of genomic complementarity and universal PCR primer sites; the LSO also contains a unique address sequence complementary to a particular bead type. The address sequence hybridizes to the universal bead type probes in the last hybridization step.

source for retrospective studies would be FFPE tissues for which patient outcomes are already known (26–28). The ability to analyze gene expression in archived tissues will greatly facilitate research that correlates gene expression profiles with given disease states or with histological and clinical phenotypes. These studies will enable the identification and validation of biomarkers useful for therapeutic decision making. Since FFPE sample collection and storage is a routine practice in pathology laboratories, any biomarkers identified from FFPE samples can be quickly adopted for future clinical test development. However, RNA extracted from FFPE tissues is often significantly degraded, which makes it hard to work with using most microarray technology (29). Because of RNA degradation, gene expression profiling in FFPE tissues has traditionally been done one gene at a time using quantitative reverse transcription PCR (RT-PCR) (28,30–35). This severely limits the number of genes that can be monitored with reasonable cost.

Illumina has developed a flexible, sensitive, accurate, and cost-effective gene expression profiling assay, the DASL™ (cDNA-mediated annealing, selection, extension and ligation) assay based on the same universal address array used with the GoldenGate SNP genotyping assay (36). The method can tolerate RNA degradation because it does not require an intact poly(A) tail for cDNA synthesis. The DASL assay multiplexes to over 1500 sequence targets (e.g., 500 genes at three probes per gene or 1500 genes at one probe per gene). Thus, it provides a mid-density expression profiling system that fills a gap between two existing RNA profiling technologies; namely, quantitative RT-PCR (high sample throughput/low gene content) and high-density oligonucleotide or cDNA microarray technology (high gene content/low sample throughput, and generally not useful with FFPE samples). The DASL assay uses 100 ng to as little as 5 ng total RNA and can resolve less than a 2-fold difference in expression between samples across a dynamic range of 2.5–3 logs. Using the DASL assay, we obtained highly reproducible tissue- and cancer-specific gene expression profiles with RNAs isolated from formalin-fixed tissues that had been stored from 1 to over 10 years (37,38) (Figure 4). Furthermore, we validated tissue- and cancer-specific markers identified from previous genome-wide expression profiling studies of fresh-frozen samples (38). We also performed a gene expression profiling study in 270 benign and prostate carcinoma FFPE tissues, including nine needle biopsy samples. We obtained highly reproducible gene expression profiles and used these to identify tumor- and stroma-specific genes and gene panels that correlated with clinical parameters such as high, medium, and low risk groups, prostate-specific antigen (PSA) velocity, relapse, sustained remission, and overall survival (Figure 5).

In summary, an automated DASL assay, coupled with a Sentrix Array Matrix (96 samples) or Sentrix BeadChip (16 samples) should prove useful for high-throughput expression profiling of hundreds or thousands of genes in large clinical FFPE sample sets (38). This technology is widely applicable to the retrospective analysis

of existing large paraffin-embedded sample collections obtained during clinical trials or from population-based cohorts. We believe this technology is especially useful for determining cancer prognosis or therapy response; instead of waiting years to accumulate fresh tissue and track patient outcomes, gene expression analysis can be performed using routinely stored tumor specimens from patients with known outcomes, therefore accelerating clinical trials. The technology is highly flexible and scalable, both in terms of the number of genes measured in a single sample and the number of samples that can be processed in parallel.

DNA METHYLATION DETECTION

DNA methylation is widespread and plays a critical role in the regulation of gene expression in development and diseases such as diabetes, schizophrenia, aging, and cancers (39–41). Changes in DNA methylation are recognized as one of the most common forms of molecular alteration in human neoplasia (42–44). Hypermethylation of CpG islands located in the promoter regions of tumor suppressor genes is now firmly established as the most frequent mechanism for gene inactivation in cancers (45,46). Most recently, gene hypermethylation was associated with clinical risk groups for neuroblastoma (47), as well as with hormone receptor status and response to tamoxifen in breast cancer (48,49).

In response to the need for high-throughput cost-effective technology for DNA methylation detection, Illumina adapted the GoldenGate SNP genotyping system to analyze bisulfite-converted genomic DNA. In this assay, nonmethylated cytosines (C) are converted to uracil (U) when treated with bisulfite, while methylated cytosines remain unchanged. Hybridization behavior of uracil is similar to that of thymine (T), allowing detection of the methylation status of a particular cytosine by carrying out a genotyping assay for a C/T polymorphism. We determine methylation status of the interrogated locus by calculation of the ratio of the fluorescent signals from the “C” (methylated) and “T” (unmethylated) alleles.

We have characterized this assay system extensively (M. Bibikova et al., manuscript submitted). We demonstrated that the current assay can detect changes in methylation status at 1536 different CpG sites simultaneously using 200 ng human genomic DNA. We used this technology to analyze the methylation status of 371 cancer-related genes (approximately 1500 CpG sites) in over 100 lung cancer (squamous cell carcinoma and adenocarcinoma) and normal tissue samples. We discovered methylation-based biomarkers for lung cancer identification, and we identified with high confidence ($P < 0.001$) panels of specific squamous cell carcinoma and adenocarcinoma methylation markers (M. Bibikova et al., manuscript submitted). These results demonstrate the effectiveness of the method for reliably profiling many CpG sites in parallel for the discovery of informative methylation patterns. The technology should prove useful for DNA methylation analysis in large populations.

Figure 4. Gene expression profiling in formalin-fixed paraffin-embedded (FFPE) samples using the DASL (cDNA-mediated annealing, selection, extension and ligation) assay. Highly reproducible gene expression profiles are obtained with RNAs isolated from the healthy and tumor FFPE tissue of the same patient (left and middle panels, respectively). Differential expression (right panel) can be readily identified between normal and cancer samples.

Figure 5. Kaplan-Meier analysis of overall survival with 47 prostate cancer samples of tumor content >10%. (A) Survival prediction based on Gleason score. (B) Survival prediction based on expression of four gene markers selected from several hundred monitored by the DASL (cDNA-mediated annealing, selection, extension and ligation) assay (M. Bibikova, J.-B. Fan, E. Chudin, et al., unpublished data).

Circulating tumor nucleic acids in blood have been demonstrated to reflect the biologic characteristics of tumors, and cancer-specific methylation signatures have been detected in plasma, serum, and other samples obtained via less invasive routes than surgical removal (50–52). A future challenge is to develop a sensitive and minimally invasive method (e.g., a blood test) to detect small amounts of cancer DNA with altered patterns of methylation in the presence of large amounts of DNA derived from normal tissues. This assay would enable the routine use of methylation patterns for early diagnosis and to classify and predict different kinds or stages of cancer, cancer therapeutic outcomes, and patient survival.

DIGITAL BEAD TECHNOLOGY

As mentioned above, Illumina is developing a comprehensive bead-based solution for biomarker R&D and in vitro and molecular diagnostic studies, including those that require low-complexity as well as high-complexity testing. For multiplex levels of 10 to several hundred, we developed a bead technology that uses a holographic imprinting approach to add a digital address to rod-shaped beads, each of which is a potential site for a biological assay. The beads can support both nucleic acid and protein probe content. This bead technology will enable easy assay customization, high sample throughput, reliable performance, and low manufacturing costs. We retain the advantage of using multiple beads of each type, and we make the beads available in a solution format to speed reactions times in a diagnostic setting.

To read out both the holographic address and the analytical signals on Illumina's digital beads, an innovative Virtual Cytometer™ scanning system has been developed. The microbeads are transferred automatically from a standard microtiter plate to a cell where they self-align into randomly ordered arrays for readout. This system is being designed to be compatible with liquid handling automation and informatics infrastructures used in high-throughput diagnostic laboratories. Thus, it will further extend the broad capability of BeadArray technology for emerging clinical research and diagnostics applications critical to the success of PGx.

CONCLUSION

Although many practical obstacles remain, there is now broad consensus that the widespread use of PGx to enable personalized medicine is no longer a question of "if," but "when." It is well established that inheritance plays a major role in drug response (53). Although efficacy and adverse reactions are often multigenic (54),

the availability of whole genome genotyping assays at low cost is eliminating the scientific barriers to unraveling complex genetic interactions. With affordable technology now available, a remaining practical barrier is the availability of well-characterized patient cohorts of sufficient size. Assays that enable use of archived FFPE tissue offer one solution. Advocates of PGx have proposed that it should be standard practice to obtain genomic DNA, with appropriate consent for PGx research, from all patients enrolled in clinical drug trials (54). Forward-looking pharmaceutical companies have accepted this challenge (55), hastening the day when complex PGx relationships will be understood and used to ensure effective, safe therapy for the many human diseases influenced by individual genetics. Since the ultimate goal of PGx is to give the right drug to the right patient (via genetic testing) with the right dose at the right time,

which requires the measurement of RNA and/or protein levels in relevant tissue(s), what is most needed is an integrated system that can provide a comprehensive molecular profile (of the patient) at genetic, epigenetic, and RNA/protein expression levels.

ACKNOWLEDGMENTS

We thank John Stuelplnagel of Illumina, and Thomas Foster and Ann Kieffaber at IMS Health for constructive suggestions and contributions to this article. We thank Marina Bibikova for use of DNA methylation data and Eugene Chudin for the analysis of unpublished data presented in Figure 5, and Frank Steemers and Kevin Gunderson for unpublished information and useful discussion about the Infinium assay. We are indebted to the many other scientists and engineers at Illumina, whose skill and dedication have created the platforms and assays reviewed here.

COMPETING INTERESTS STATEMENT

The authors are employed, or were employed, by Illumina, Inc., which has developed the products and assays described in this article.

REFERENCES

1. Frueh, F.W., F. Goodsaid, A. Rudman, S.M. Huang, and L.J. Lesko. 2005. The need for education in pharmacogenomics: a regulatory perspective. *Pharmacogenomics J.* 5:218-220.
2. Lesko, L.J. and J. Woodcock. 2004. Translation of pharmacogenomics and pharmacogenetics: a regulatory perspective. *Nat. Rev. Drug Discov.* 3:763-769.
3. The International HapMap Consortium. 2003. The International HapMap Project. *Nature* 426:789-796.
4. Oliphant, A., D.L. Barker, J.R. Stuelplnagel, and M.S. Chee. 2002. BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. *BioTechniques Suppl.* 56-61.
5. Gunderson, K.L., S. Kruglyak, M.S. Graige, F. Garcia, B.G. Kermani, C. Zhao, D. Che, T. Dickinson, et al. 2004. Decoding randomly ordered DNA arrays. *Genome Res.* 14:870-877.
6. Barker, D.L., D. Theriault, D. Che, T. Dickinson, R. Shen, and R. Kain. 2003. Self-assembled random arrays: High-performance imaging and genomics applications on a high-density microarray platform. *Proc. SPIE* 4966:1-11.
7. Gunderson, K.L., F.J. Steemers, G. Lee, L.G. Mendoza, and M.S. Chee. 2005. A genome-wide scalable SNP genotyping assay using microarray technology. *Nat. Genet.* 37:549-554.
8. Fan, J.B., A. Oliphant, R. Shen, B.G. Kermani, F. Garcia, K.L. Gunderson, M. Hansen, F. Steemers, et al. 2003. Highly parallel SNP genotyping. *Cold Spring Harb. Symp. Quant. Biol.* 68:69-78.
9. Shen, R., J.B. Fan, D. Campbell, W. Chang, J. Chen, D. Doucet, J. Yeakley, M. Bibikova, et al. 2005. High-throughput SNP genotyping on universal bead arrays. *Mutat. Res.* 573:70-82.
10. Barker, D.L., M.S. Hansen, A.F. Faruqi, D. Giannola, O.R. Irsula, R.S. Lasken, M. Latterich, V. Makarov, et al. 2004. Two methods of whole-genome amplification enable accurate genotyping across a 2320-SNP linkage panel. *Genome Res.* 14:901-907.

11. Golub, T.R., D.K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J.P. Mesirov, H. Coller, M.L. Loh, et al. 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286:531-537.
12. Perou, C.M., T. Sorlie, M.B. Eisen, M. van de Rijn, S.S. Jeffrey, C.A. Rees, J.R. Pollack, D.T. Ross, et al. 2000. Molecular portraits of human breast tumours. *Nature* 406:747-752.
13. Welsh, J.B., L.M. Sapinoso, A.I. Su, S.G. Kern, J. Wang-Rodriguez, C.A. Moskaluk, H.F. Frierson, Jr., and G.M. Hampton. 2001. Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. *Cancer Res.* 61:5974-5978.
14. Dhanasekaran, S.M., T.R. Barrette, D. Ghosh, R. Shah, S. Varambally, K. Kurachi, K.J. Pienta, M.A. Rubin, et al. 2001. Delineation of prognostic biomarkers in prostate cancer. *Nature* 412:822-826.
15. West, M., C. Blanchette, H. Dressman, E. Huang, S. Ishida, R. Spang, H. Zuzan, J.A. Olson, Jr., et al. 2001. Predicting the clinical status of human breast cancer by using gene expression profiles. *Proc. Natl. Acad. Sci. USA* 98:11462-11467.
16. van 't Veer, L.J., H. Dai, M.J. van de Vijver, Y.D. He, A.A. Hart, M. Mao, H.L. Peterse, K. van der Kooy, et al. 2002. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415:530-536.
17. Welsh, J.B., L.M. Sapinoso, S.G. Kern, D.A. Brown, T. Liu, A.R. Bauskin, R.L. Ward, N.J. Hawkins, et al. 2003. Large-scale delineation of secreted protein biomarkers overexpressed in cancer tissue and serum. *Proc. Natl. Acad. Sci. USA* 100:3410-3415.
18. Chuaiqui, R.F., R.F. Bonner, C.J. Best, J.W. Gillespie, M.J. Flaig, S.M. Hewitt, J.L. Phillips, D.B. Krizman, et al. 2002. Post-analysis follow-up and validation of microarray experiments. *Nat. Genet.* 32(Suppl):509-514.
19. van de Vijver, M.J., Y.D. He, L.J. van't Veer, H. Dai, A.A. Hart, D.W. Voskuil, G.J. Schreiber, J.L. Peterse, et al. 2002. A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* 347:1999-2009.
20. Stuart, R.O., W. Wachsmen, C.C. Berry, J. Wang-Rodriguez, L. Wasserman, I. Kiacansky, D. Masys, K. Arden, et al. 2004. In silico dissection of cell-type-associated patterns of gene expression in prostate cancer. *Proc. Natl. Acad. Sci. USA* 101:615-620.
21. Lossos, I.S., D.K. Czerwinski, A.A. Alizadeh, M.A. Wechsler, R. Tibshirani, D. Botstein, and R. Levy. 2004. Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. *N. Engl. J. Med.* 350:1828-1837.
22. Bullinger, L., K. Dohner, E. Bair, S. Frohling, R.F. Schlenk, R. Tibshirani, H. Dohner, and J.R. Pollack. 2004. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N. Engl. J. Med.* 350:1605-1616.
23. Ramaswamy, S. 2004. Translating cancer genomics into clinical oncology. *N. Engl. J. Med.* 350:1814-1816.
24. Paik, S., S. Shak, G. Tang, C. Kim, J. Baker, M. Cronin, F.L. Baehner, M.G. Walker, et al. 2004. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med.* 351:2817-2826.
25. Wang, Y., J.G. Klijn, Y. Zhang, A.M. Sieuwerts, M.P. Look, F. Yang, D. Talantov, M. Timmermans, et al. 2005. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 365:671-679.
26. Lewis, F., N.J. Maughan, V. Smith, K. Hillan, and P. Quirke. 2001. Unlocking the archive—gene expression in paraffin-embedded tissue. *J. Pathol.* 195:66-71.
27. Ramaswamy, S. and T.R. Golub. 2002. DNA microarrays in clinical oncology. *J. Clin. Oncol.* 20:1932-1941.
28. Abrahamsen, H.N., T. Steiniche, E. Nexø, S.J. Hamilton-Dutoit, and B.S. Sorensen. 2003. Towards quantitative mRNA analysis in paraffin-embedded tissues using real-time reverse transcriptase-polymerase chain reaction: a methodological study on lymph nodes from melanoma patients. *J. Mol. Diagn.* 5:34-41.
29. Karsten, S.L., V.M. van Deerlin, C. Sabatti, L.H. Gill, and D.H. Geschwind. 2002. An evaluation of tyramide signal amplification and archived fixed and frozen tissue in microarray gene expression analysis. *Nucleic Acids Res.* 30:E4.
30. Daigo, Y., S.F. Chin, K.L. Gorringer, L.G. Bobrow, B.A. Ponder, P.D. Pharoah, and C. Caldas. 2001. Degenerate oligonucleotide primed-polymerase chain reaction-based array comparative genomic hybridization for extensive amplicon profiling of breast cancers: a new approach for the molecular analysis of paraffin-embedded cancer tissue. *Am. J. Pathol.* 158:1623-1631.
31. Godfrey, T.E., S.H. Kim, M. Chavira, D.W. Ruff, R.S. Warren, J.W. Gray, and R.H. Jensen. 2000. Quantitative mRNA expression analysis from formalin-fixed, paraffin-embedded tissues using 5' nuclease quantitative reverse transcription-polymerase chain reaction. *J. Mol. Diagn.* 2:84-91.
32. Lehmann, U., S. Glockner, W. Kleeberger, H.F. von Wasielewski, and H. Kreipe. 2000. Detection of gene amplification in archival breast cancer specimens by laser-assisted microdissection and quantitative real-time polymerase chain reaction. *Am. J. Pathol.* 156:1855-1864.
33. Palmieri, G., P.A. Ascierto, A. Cossu, N. Mozzillo, M.L. Motti, S.M. Satriano, G. Botti, C. Caraco, et al. 2001. Detection of occult melanoma cells in paraffin-embedded histologically negative sentinel lymph nodes using a reverse transcriptase polymerase chain reaction assay. *J. Clin. Oncol.* 19:1437-1443.
34. Specht, K., T. Richter, U. Muller, A. Walch, M. Werner, and H. Hofler. 2001. Quantitative gene expression analysis in microdissected archival formalin-fixed and paraffin-embedded tumor tissue. *Am. J. Pathol.* 158:419-429.
35. Cronin, M., M. Pho, D. Dutta, J.C. Stephens, S. Shak, M.C. Kiefer, J.M. Esteban, and J.B. Baker. 2004. Measurement of gene expression in archival paraffin-embedded tissues: development and performance of a 92-gene reverse transcriptase-polymerase chain reaction assay. *Am. J. Pathol.* 164:35-42.
36. Fan, J.B., J.M. Yeakley, M. Bibikova, E. Chudin, E. Wickham, J. Chen, D. Doucet, P. Rigault, et al. 2004. A versatile assay for high-throughput gene expression profiling on universal array matrices. *Genome Res.* 14:878-885.
37. Bibikova, M., J.M. Yeakley, E. Chudin, J. Chen, E. Wickham, J. Wang-Rodriguez, and J.B. Fan. 2004. Gene expression profiles in formalin-fixed, paraffin-embedded tissues obtained with a novel assay for microarray analysis. *Clin. Chem.* 50:2384-2386.
38. Bibikova, M., D. Talantov, E. Chudin, J.M. Yeakley, J. Chen, D. Doucet, E. Wickham, D. Atkins, et al. 2004. Quantitative gene expression profiling in formalin-fixed, paraffin-embedded tissues using universal bead arrays. *Am. J. Pathol.* 165:1799-1807.
39. Egger, G., G. Liang, A. Aparicio, and P.A. Jones. 2004. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* 429:457-463.
40. Laird, P.W. and R. Jaenisch. 1996. The role of DNA methylation in cancer genetic and epigenetics. *Annu. Rev. Genet.* 30:441-464.
41. Li, E., C. Beard, and R. Jaenisch. 1993. Role for DNA methylation in genomic imprinting. *Nature* 366:362-365.
42. Balmain, A., J. Gray, and B. Ponder. 2003. The genetics and genomics of cancer. *Nat. Genet.* 33(Suppl):238-244.
43. Baylin, S.B. and J.G. Herman. 2000. DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends Genet.* 16:168-174.
44. Feinberg, A.P. and B. Tycko. 2004. The history of cancer epigenetics. *Nat. Rev. Cancer* 4:143-153.
45. Esteller, M. 2002. CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene* 21:5427-5440.
46. Herman, J.G. and S.B. Baylin. 2003. Gene silencing in cancer in association with promoter hypermethylation. *N. Engl. J. Med.* 349:2042-2054.
47. Alaminos, M., V. Davalos, N.K. Cheung, W.L. Gerald, and M. Esteller. 2004. Clustering of gene hypermethylation associated with clinical risk groups in neuroblastoma. *J. Natl. Cancer Inst.* 96:1208-1219.
48. Martens, J.W., I. Nimmrich, T. Koenig, M.P. Look, N. Harbeck, F. Model, A. Kluth, J. Bolt-de Vries, et al. 2005. Association of DNA methylation of phosphoserine aminotransferase with response to endocrine therapy in patients with recurrent breast cancer. *Cancer Res.* 65:4101-4117.
49. Widschwendter, M., K.D. Siegmund, H.M. Muller, H. Fiegl, C. Marth, E. Muller-Holzner, P.A. Jones, and P.W. Laird. 2004. Association of breast cancer DNA methylation profiles with hormone receptor status and response to tamoxifen. *Cancer Res.* 64:3807-3813.
50. Esteller, M., M. Sanchez-Cespedes, R. Rosell, D. Sidransky, S.B. Baylin, and J.G. Herman. 1999. Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. *Cancer Res.* 59:67-70.
51. Herman, J.G. 2004. Circulating methylated DNA. *Ann. NY Acad. Sci.* 1022:33-39.
52. Muller, H.M., H. Fiegl, A. Widschwendter, and M. Widschwendter. 2004. Prognostic DNA methylation marker in serum of cancer patients. *Ann. NY Acad. Sci.* 1022:44-49.
53. Weinshilboum, R. 2003. Inheritance and drug response. *N. Engl. J. Med.* 348:529-537.
54. Evans, W.E. and H.L. McLeod. 2003. Pharmacogenomics—drug disposition, drug targets, and side effects. *N. Engl. J. Med.* 348:538-549.
55. Roses, A.D. 2002. Genome-based pharmacogenetics and the pharmaceutical industry. *Nat. Rev. Drug Discov.* 1:541-549.

Address correspondence to:

David L. Barker
 Illumina, Inc.
 9885 Towne Centre Drive
 San Diego, CA 92121, USA
 e-mail: dbarker@illumina.com