Five-Year, Multi-Institution Validation Studies of Archived Amplified Genomic DNA, Methylated DNA, and cDNA: Enhancing Tissue Archives With Nucleic Acid Archives

J. Langmore, V. Makarov, E. Kamberov, T. Kurihara, E. Bruening, T. Tesmer, J. M'Mwirichia, Rubicon Genomics, 4370 Varsity Drive, Ann Arbor, MI 48108, (734) 368-1705 langmore@rubicongenomics.com www.rubicongenomics.com

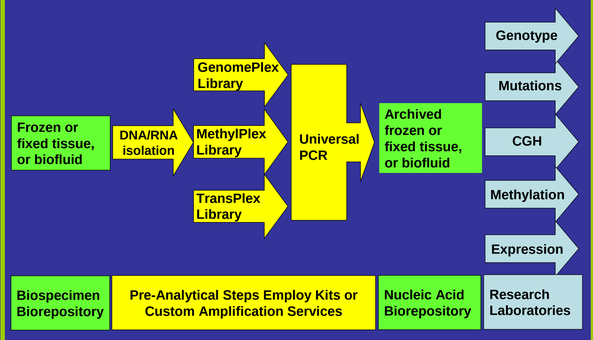
ABSTRACT

Epidemiologic, pharmacologic, and diagnostic studies link clinical with molecular data via analysis of the nucleic acids, proteins, and metabolites in patient samples. Although clinical data are immortal, clinical specimens are not. While the number and power of molecular studies grow exponentially, the numbers of clinical specimens grow very slowly due to the expense of recruiting and following up patients as well as depletion and degradation of existing samples. Typically 1 - 4 mL of serum or plasma, less than 200 mL of urine and between 1 – 10 mm of diseased tissue are archived from each patient, and have a shelf-life of less than 20 years. The quantity and quality of investigator research is severely compromised by the great difficulty in accessing the limited quantities of patient material, high cost and irreproducibility of cutting, staining, characterizing, and macro dissecting tissue samples, or of purifying molecules from individual sections or fluid aliquots.

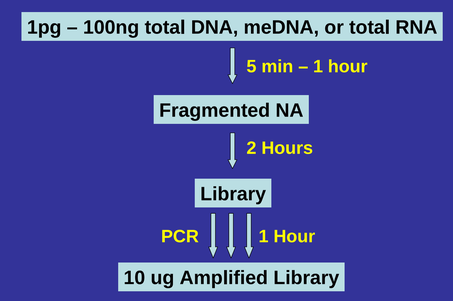
Greater utilization of archived samples for genetic, epigenetic, and expression studies can be achieved by amplifying, aliquoting, and storing nucleic acids from small portion of each patient sample. Isolation, 10,000-fold amplification, and QC of the DNA, methylated DNA or cDNA costs only ~$1 per microgram of amplified product—a small fraction of the cost of recruiting the patient and clinical data. Aliquots of amplified DNA are stored without degradation and can be shared with thousands of researchers for over 100 years, thus extending and increasing the long-term value of the patient sample and data. After DNA amplification, individual biofluid aliquots or microtome sections can be used repeatedly and reproducibly by different researchers doing complementary genetic, epigenetic, and expression studies using qPCR, microarrays and next-generation sequencing. These amplified nucleic acid archives are needed to meet the increasing demands of these high throughput analytical platforms

Published and unpublished data from over five years of experience in pharmaceutical, diagnostic, university, and government laboratories using more than 30,000 frozen and fixed tissue, serum, plasma, and urine samples will be presented to document the precision, accuracy, clinical significance, and economics of analyzing amplified DNA, meDNA, and cDNA from cancer and normal patients.

Archived Biospecimens Can Be Amplified as DNA, Methylated DNA, and cDNA to Create Inexhaustable Archives of Nucleic Acids



Amplification of Frozen or Fixed Tissue, or Biofluid is a Robust, Automatable 3-Step Process



1. Amplification works equally well for intact or highly degraded nucleic acids.
2. Nucleic acid can be reamplified a billion-fold without loss of representation
3. Extremely low background allows single cell and single chromosome amplification.

Practical Advantages of Nucleic Acid Archives

- Only a small fraction of the biospecimen needs to be extracted and amplified to produce DNA and RNA for hundreds of investigators.
- Microtomy, dissection, DNA/RNA extraction, and quality control are needed only once per biosample.
- Archived aliquots of amplified DNA, meDNA, and cDNA can be retrieved and supplied to investigators within one day and for as little as $1/microgram.
- Exactly the same DNA or cDNA sample can be studied by different investigators at different times.
- Archived amplified DNA, meDNA, and cDNA stable for decades, without change in signal or background.

GenomePlex Whole Genome Amplification

Research kits sold by Sigma-Aldrich
Services and diagnostic kits sold by Rubicon

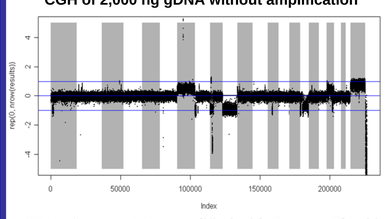
Archiving of GenomePlex WGA DNA

- Library synthesis takes ~ 2 h; 1,000X amplification takes 1 h.
- Amplified DNA performs as well as gDNA for genotyping, CGH, and sequencing.
- Low background allows genetic studies of single cells and chromosomes.
- GenomePlex amplifies highly degraded DNA (e.g., FFPE, serum/plasma, urine) for all applications except genotyping with long STRs.
- Excellent array profiling possible from 20 ng unfixed DNA or 200 ng fixed DNA.
- WGA DNA is stable if stored at -20 C, and can be reamplified >1E09 times without significant loss of representation—even for very GC-rich sequences.
- A 20 ng aliquot of biospecimen DNA can be amplified to >20 mg to supply DNA to thousands of investigators
- GenomePlex in current use to discover disease genes and biomarkers, to manufacture MDx products, and for diagnosis of human disease.

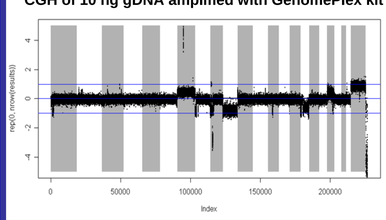
THE CANCER GENOME ATLAS STUDY OF AMPLIFIED DNA ON CGH ARRAYS

Analysis of genomic copy number alterations using Agilent oligo 244K microarray and WGA. Below is a comparison of two aCGH profiles of the same GBM sample. Top profile has been obtained by a standard protocol required 2 mkg gDNA. Lower profile is the result of GenomePlex WGA from 10 ng gDNA. Data from Dr. Alexei Protopopov of the Dana Farber Cancer Institute. (Published Sept. 2007 in TCGA Newsletter)

CGH of 2,000 ng gDNA without amplification

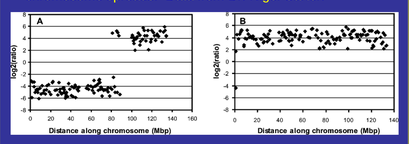


CGH of 10 ng gDNA amplified with GenomePlex kit



SANGER INSTITUTE CGH STUDY OF SINGLE COPYS OF SORTED CHROMOSOMES

Single human chromosomes were sorted into wells of microplates. The wells were amplified at Rubicon using a single-cell GenomePlex kit. Amplified DNA from single wells were labeled and hybridized to BAC microarrays and as FISH probes. Translocations were mapped on the microarrays and chromosome spreads. Data from Dr. Nigel Carter.



Refs from Nigel Carter (Sanger Institute)

1. Gribble et al. (2004) Applications of Combined DNA Microarray and Chromosome Sorting Technologies. Chromosome Research 12:35-43.
2. Gribble et al. (2004) Chromosome Paints from Single Copy Chromosomes. Chromosome Research 12: 143–151.
3. Fieger et al. High Resolution Array CGH of Single Cells. (2006) NAR.

CHROMATIN IMMUNOPRECIPITATION STUDY OF THE FIDELITY OF GENOMEPLEX AMPLIFICATION AND LINKER-MEDIATED PCR (Farnham lab)—

Chromatin immunoprecipitation (ChIP) has proven to be a powerful tool, allowing the detection of protein-DNA interactions in living cells. A single reaction does not yield enough DNA to perform genomic profiling of protein binding, so amplification is necessary. GenomePlex was found to accurately represent the ChIP DNA much better than LM-PCR (Peggy Farnham laboratory)

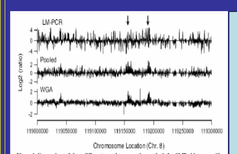


Figure 1. Comparison of three different sample preparation methods for ChIP-chip assays. The hybridization profile of a 300-kb region of chromosome 8 surrounding the EXT1 gene is shown for samples prepared by the LM-PCR (top panel), pooled (middle panel), or WGA (bottom panel) methods. Oct4 binding sites confirmed by PCR are indicated with arrows. ChIP, chromatin immunoprecipitation; LM-PCR, ligation-mediated PCR; WGA, whole genome amplification.

Table 1. Comparison of Sample Preparation Methods

Method	Total Peaks[a]	Overlapping[b]	Overlap (%)
LM-PCR	543	82	15
Pooled	491	343	70
WGA	449	280	63

LM-PCR, ligation-mediated PCR; WGA, whole genome amplification.
[a]Total number of peaks called on both arrays.
[b]If at least one of the ends of a peak region from one array overlapped a peak region from the other array, the peaks were considered to be overlapping.

1. O'Geen et al. (2006) Comparison of Sample Preparation Methods for ChIP-chip Assays. BioTechniques 41:577-580.
2. Acevedo, et al. (2007) Genome-Scale ChIP-chip Analysis Using 10,000 Human Cells. BioTechniques 43:791-797.

ILLUMINA AND GSK STUDY OF GENOTYPING ACCURACY OF GENOMEPLEX WGA—

Genomic and WGA DNA were genotyped using the Illumina BeadArray™. Ten samples were analyzed at 2,320 sites. All data were pooled as there was no significant difference in number of calls or concordance among samples. The high concordance between gDNA and WGA DNA results demonstrate the accuracy and reproducibility of GenomePlex.

Analysis of Genotyping Results		
	Genomic DNA	Amplified DNA
Assays Run	23,200	23,200
Calls Made	23,159	23,200
% Total	99.8%	100%
% Genomic Concordance*	-	99.8%

1. Barker et al. (2004) Two Methods of Whole-Genome Amplification Enable Accurate Genotyping Across a 2320 SNP Linkage Panel. Genome Research 901-907.
2. Barnes et al., (2006) Polymorphisms in the novel gene acyloxyacyl hydrogenase (AOAH) are associated with asthma and associated phenotypes. J. Allergy and Clinical Immunology 118:70-77.
3. Barnes et al. (2006) Variants in the gene encoding C3 are associated with asthma and related phenotypes among African Caribbean families. Genes and Immunity 7:27-35.

USE OF GENOMEPLEX FOR ARRAY CGH OF MICRODISSECTED FFPE SECTIONS (Du lab, Cambridge UK.

Johnson et al. (2006) Application of Array CGH on Archival Formalin-Fixed Paraffin-Embedded Tissues Including Small Numbers of Microdissected Cells. Laboratory Investigation (2006) 1 – 11.

Table 2 aCGH using DNA from FFPE tissue: minimum requirement for DNA integrity, DNA quantity or number of cells

Sample	Maximum DNA fragment (bp) amplified	Quality of aCGH by visual inspection	Pearson's correlation to aCGH data from corresponding frozen tissue		
			400 ng DNA from FFPE tissue	20 ng DNA from FFPE tissue with WGA	2000 cells from FFPE tissue with WGA
Xenografts					
X1	400	Good	0.98	0.96	0.91
X2	400	Poor	0.51	0.90	0.91
X3	400	Good	0.88	0.88	0.93
X4	400	Good	0.96	0.89	
X5	300	Good	0.83	0.82	
X6	100	Poor	0.72		
X7	100	Poor	0.77		
X8	100	Poor	0.50		
X9	100	Poor	0.51		
X10	100	Failed to label			
X11	100	Failed to label			
Primary glioblastoma					
P1	400	Good	0.91		
P2	400	Good	0.87		0.91
P3	300	Good	0.86		
P4	300	Good	0.91		

[a]Poor correlation was due to the presence of >20% necrosis.

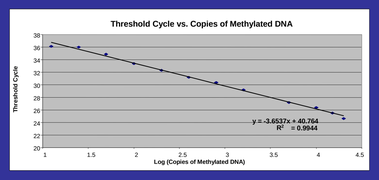
MethylPlex Whole Methylome Amplification

Diagnostic Partnerships and Beta Kits Available from Rubicon

Archiving of MethylPlex WMA DNA

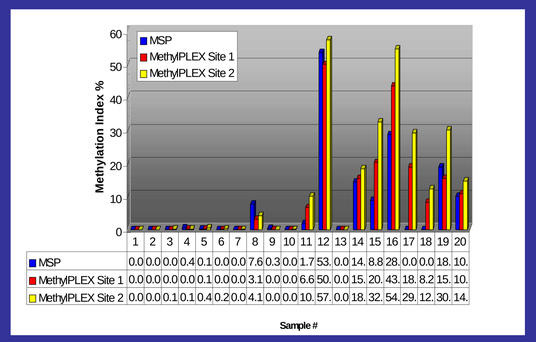
- Library synthesis takes ~ 3 h; 1,000X amplification takes 1 h.
- NO bisulfite converstion; no methylation-specific PCR assays to design; no custom arrays necessary.
- Same clinical sensitivity and specificity as best bisulfite-based tests.
- 2 mL of plasma
- Amplified MethylPlex DNA used to assay for DNA methylation using any copy number assay, such as qPCR, promoter and CpG island arrays, or Next-Generation sequencing platforms.
- Low background allows methylation studies of as little as 10 cells, and as little as 0.01% methylation
- MethylPlex works on highly degraded DNA (e.g., FFPE, serum/plasma, urine) for all methylation assays.
- Excellent array methylation profiling from 20 ng unfixed DNA or 200 ng fixed DNA.
- MethylPlex WMA DNA is stable if stored at -20 C, and can be reamplified >1E06 times without significant loss of sensitivity.
- A 20 ng aliquot of biospecimen DNA can be amplified to >20 mg of MethylPlex DNA to enable thousands of investigators to study methylation in any part of the genome
- MethylPlex has been used to discover biomarkers, and develop patient tests.

MethylPlex real-time PCR assays can be used for sensitive detection of trace amounts of methylated cell-free DNA in urine. Artificially-methylated urine DNA was combined with unmethylated urine DNA at different ratios, and amplified methylome libraries were prepared from those samples. Quantitative real-time PCR analysis for methylation at Gene#5 promoter region demonstrated linear detection of methylation down to 10 cell equivalents, or 0.1% of the total.



MethylPlex and Bisulfite Test Concordance

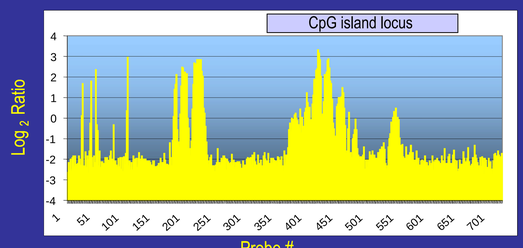
MethylPlex Q-PCR assays are highly concordant with MethyLight assays for the same promoters. In this example of the promoter for APC1, the methylation indexes of the bisulfite-based and 2 MethylPlex assays were compared for 20 normal and twenty prostate cancer tissue samples. The analytical and clinical performance of both types of assays were very similar.



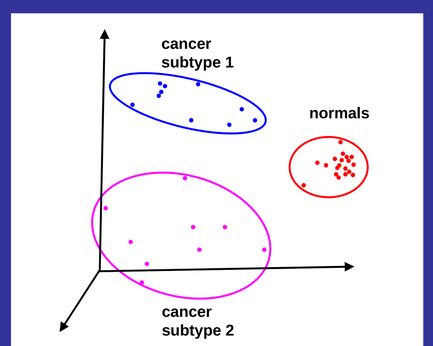
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
MSP	0.0	0.0	0.0	0.0	0.4	0.1	0.0	0.0	0.0	7.6	0.3	0.0	1.7	53.	0.0	0.0	14.	8.8	28.	0.0
MethylPLEX Site 1	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	3.1	0.0	0.0	0.0	6.6	50.	0.0	15.	20.	43.	18.	8.2
MethylPLEX Site 2	0.0	0.0	0.0	0.1	0.1	0.4	0.2	0.0	4.1	0.0	0.0	0.0	10.	57.	0.0	18.	32.	54.	29.	12.

	18	19	20
MSP	0.0	18.	10.
MethylPLEX Site 1	15.	10.	
MethylPLEX Site 2	30.	14.	

METHYLPLEX CAN BE USED TO PROFILE METHYLATION ACROSS THE ENTIRE GENOME USING AGILENT OR NIMBLEGEN PROMOTER ARRAYS. This slide shows ratio of methylation in a cancer tissue divided by the methylation of a normal tissue on an oligonucleotide array region with 5 b tiling. The differential methylation is detected across one of the CpG islands.



METHYLPLEX CAN BE USED TO DISCOVER BIOMARKERS THAT CAN DISTINGUISH TWO SUBTYPES OF THIS CANCER FROM NORMAL, ADJACENT TISSUE FROM THE SAME PATIENTS. This example is from a study of FFPE sections from the tumors and adjacent normal tissue.



TransPlex Whole Transcriptome Amplification

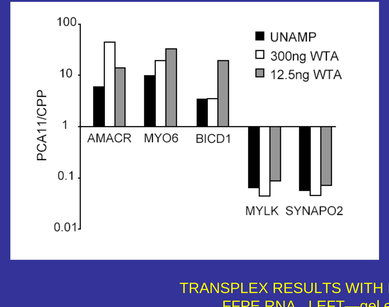
Research kits sold by Sigma-Aldrich
Services and diagnostic kits sold by Rubicon

Archiving of TransPlex WTA cDNA

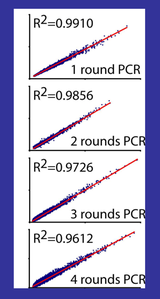
- Library synthesis takes ~ 2 h; 1,000X amplification takes 1 h.
- Much faster and simpler than IVT.
- Amplified cDNA performs as well as unamplified cDNA for qPCR assays or profiling of gene or exon expression.
- No 3' bias or exon bias, as found after IVT.
- TransPlex amplifies highly degraded RNA (e.g., FFPE, serum/plasma, urine) for all applications.
- Excellent array profiling possible from 20 ng unfixed RNA or 200 ng fixed RNA.
- WTA cDNA is stable if stored at -20 C, and can be reamplified >1E09 times without significant loss of representation.
- A 20 ng aliquot of biospecimen RNA can be amplified to >20 mg to supply DNA to thousands of investigators
- TransPlex has been used to discover disease genes and biomarkers.

CONCORDANCE BETWEEN Q-PCR EXPRESSION OF GENES BEFORE AND AFTER TRANSPLEX AMPLIFICATION (Chinniayan lab)

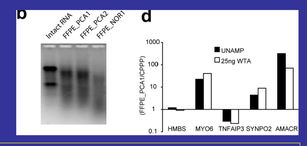
single amplification of 12 or 300 ng RNA



Two aliquots from the same sample, amplified different numbers of times and hybridized against each other, showing excellent reproducibility



TRANSPLEX RESULTS WITH HIGHLY DEGRADED FFPE RNA. LEFT—gel electrophoretic pattern RIGHT—Q-PCR results from FFPE sample before and after amplification. Amplification does not change representation



NSABP GENE EXPRESSION PROFILING OF B-27.2 PRE-TREATMENT FFPE CORE BIOPSY SECTIONS TO PREDICT RISK OF BREAST CANCER RECURRENCE (Dr. Soon Paik, recent update to NCI)—

Study of 326 core-biopsy cases show that expression of ESR1 has much better prognostic value than pathological complete response, node status, or TRT.

STUDY DESIGN—

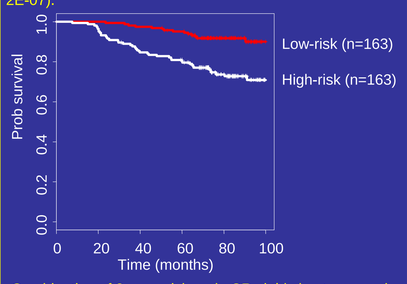
•100 ng total RNA as starting material

•RNA amplified to 20 µg using Rubicon TransPlex custom amplification service

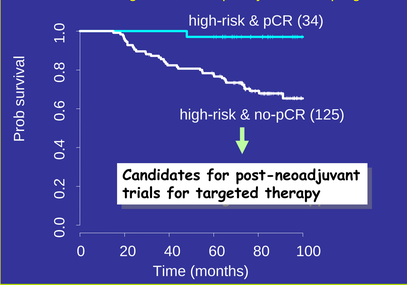
•Hybridization to Affymetrix GeneChip U133 2.0 plus

•PAM and SUPERPC used for prediction of ER, pCR, and outcome. Microarray data was 96% concordant with IHC. ESR1, BCL2, and GATA3 were best predictors.

RESULTS—All FFPE core biopsies gave interpretable data. Three-gene test predicted survival with high significance (p< 2E-07).



Combination of 3-gene risk and pCR yields best prognosis



GENE EXPRESSION PROFILING FROM ARCHIVED FROZEN AND FIXED PROSTATE TISSUE AND URINE (Arul Chinniayan, Univ. Michigan)

Seven studies to study expression of genes involved in prostate cancer progression and diagnosis.

STUDY DESIGN—

•Studies use WTA cDNA from frozen and fixed tissue, urine.

•Array studies employ as few as 2,000 cells.

•cDNA and oligonucleotide arrays used.

•Expression results correlated to gene fusion and amplification events studied by WGA of the same samples.

RESULTS—Genes and pathways most correlated with prostate cancer identified, as well as identification of urine expression biomarkers.

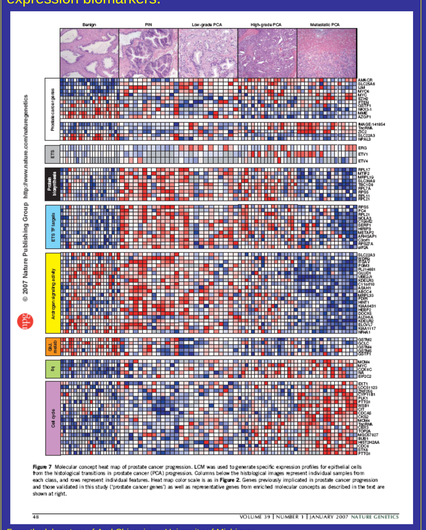


Figure 7 Molecular concept heat map of prostate cancer progression. LCM was used to generate specific expression profiles for epithelial cells from the histological transitions in prostate cancer (PCa) progression. Columns below the histological images represent individual samples from each class, and rows represent individual features. Heat map color scale is as in Figure 2. Genes previously implicated in prostate cancer progression and those validated in this study (prostate cancer genes) as well as representative genes from enriched molecular concepts as described in the text are shown at right.

From the laboratory of Arul Chinnaiyan, University of Michigan:
1. Tomlins et al. (2006). Whole Transcriptome Amplification for Gene expression Profiling and Development of Molecular Archives. Neoplasia, 8: 153-162.
2. Laxman et al. (2006) Noninvasive Detection of TMPRSS2:ERG Fusion Transcripts in the Urine of Men with Prostate Cancer. Neoplasia, 8:885-888.
3. Tomlins et al. (2007). Integrative Molecular Concept Modeling of Prostate Cancer Progression. Nature Genetics, 39:41-51.
4. Kim et al. (2007) Integrative Analysis of Genomic Aberrations Associated with Prostate Cancer Progression. Cancer Research 67:8229-8239.
5. Laxman et al. (2008). A First-generation Multiplex Biomarker Analysis of Urine for the Early Diagnosis of Prostate Cancer. Cancer Research, 68:645-649.